(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
19 June 2003 (19.06.2003)

PCT

(10) **International Publication Number**
WO 03/050084 A2

(51) **International Patent Classification⁷:** C07D

(21) **International Application Number:** PCT/US02/39126

(22) **International Filing Date:** 6 December 2002 (06.12.2002)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**

60/340,235	7 December 2001 (07.12.2001)	US
60/343,007	19 December 2001 (19.12.2001)	US
60/343,546	21 December 2001 (21.12.2001)	US
60/354,388	4 February 2002 (04.02.2002)	US
60/357,675	15 February 2002 (15.02.2002)	US

(71) **Applicant** *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **KABLE, Amy, E.** [US/US]; 2345 Polk Street #4, San Francisco, CA 94109 (US). **CHIEN, David** [US/US]; 444 Yerba Buena Avenue, Los Altos, CA 94022 (US). **WILSON, Amy, D.** [US/US]; 1056 Continentals Way #27, Belmont, CA 94002 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **GORVAD, Ann, E.** [US/US]; 369 Marie Common, Livermore, CA 94550 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **EMERLING, Brooke, M.** [US/US]; 1735 Woodland Avenue # 71, Palo Alto, CA 94303 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **JIN, Pei** [US/US]; 320 Curtner Avenue #D, Palo Alto, CA 94306 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square #712, Union City, CA 94587 (US). **LEHR-MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US). **KHARE, Reena** [US/US]; 12650 Orella Court, Saratoga, CA 95070 (US). **LEE, Sally** [US/US]; 3643 26th Street, San Francisco, CA 94110 (US). **HAWKINS, Phillip, R.** [US/US]; 750 North Shorline Boulevard #115, Mountain View, CA 94043 (US). **BECHA, Shanya, D.** [US/US]; 21062 Gary Drive # 117, Castro Valley, CA 94546 (US). **LEE, Soo, Yeun** [KR/US]; 40 Westdale Avenue, Daly City, CA 94015 (US). **SPRAGUE, William, W.** [US/US]; 611 13th Street # C, Sacramento, CA 95814 (US). **ZEBARJADIAN, Yeganeh** [IR/US]; 830 Junipero Serra Boulevard, San Francisco, CA 94127 (US).

(74) **Agents: HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** KINASES AND PHOSPHATASES

(57) **Abstract:** Various embodiments of the invention provide human kinases and phosphatases (KPP) and polynucleotides which identify and encode KPP. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing, treating, or preventing disorders associated with aberrant expression of KPP.

WO 03/050084 A2

KINASES AND PHOSPHATASES

TECHNICAL FIELD

The invention relates to novel nucleic acids, kinases and phosphatases encoded by these nucleic acids, and to the use of these nucleic acids and proteins in the diagnosis, treatment, and prevention of cardiovascular diseases, immune system disorders, neurological disorders, disorders affecting growth and development, lipid disorders, cell proliferative disorders, and cancers. The invention also relates to the assessment of the effects of exogenous compounds on the expression of nucleic acids and kinases and phosphatases.

BACKGROUND OF THE INVENTION

Reversible protein phosphorylation is the ubiquitous strategy used to control many of the intracellular events in eukaryotic cells. It is estimated that more than ten percent of proteins active in a typical mammalian cell are phosphorylated. Kinases catalyze the transfer of high-energy phosphate groups from adenosine triphosphate (ATP) to target proteins on the hydroxyamino acid residues serine, threonine, or tyrosine. Phosphatases, in contrast, remove these phosphate groups. Extracellular signals including hormones, neurotransmitters, and growth and differentiation factors can activate kinases, which can occur as cell surface receptors or as the activator of the final effector protein, as well as other locations along the signal transduction pathway. Cascades of kinases occur, as well as kinases sensitive to second messenger molecules. This system allows for the amplification of weak signals (low abundance growth factor molecules, for example), as well as the synthesis of many weak signals into an all-or-nothing response. Phosphatases, then, are essential in determining the extent of phosphorylation in the cell and, together with kinases, regulate key cellular processes such as metabolic enzyme activity, proliferation, cell growth and differentiation, cell adhesion, and cell cycle progression.

KINASES

Kinases comprise the largest known enzyme superfamily and vary widely in their target molecules. Kinases catalyze the transfer of high energy phosphate groups from a phosphate donor to a phosphate acceptor. Nucleotides usually serve as the phosphate donor in these reactions, with most kinases utilizing adenosine triphosphate (ATP). The phosphate acceptor can be any of a variety of molecules, including nucleosides, nucleotides, lipids, carbohydrates, and proteins. Proteins are phosphorylated on hydroxyamino acids. Addition of a phosphate group alters the local charge on the acceptor molecule, causing internal conformational changes and potentially influencing intermolecular contacts. Reversible protein phosphorylation is the primary method for regulating protein activity in eukaryotic cells. In general, proteins are activated by phosphorylation in response to extracellular signals such as hormones, neurotransmitters, and growth and differentiation factors. The activated proteins initiate the cell's intracellular response by way of intracellular signaling pathways and second messenger molecules such as cyclic nucleotides, calcium-calmodulin, inositol, and various mitogens, that regulate protein phosphorylation.

Kinases are involved in all aspects of a cell's function, from basic metabolic processes, such as glycolysis, to cell-cycle regulation, differentiation, and communication with the extracellular environment through signal transduction cascades. Inappropriate phosphorylation of proteins in cells has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

There are two classes of protein kinases. One class, protein tyrosine kinases (PTKs), phosphorylates tyrosine residues, and the other class, protein serine/threonine kinases (STKs), phosphorylates serine and threonine residues. Some PTKs and STKs possess structural characteristics of both families and have dual specificity for both tyrosine and serine/threonine residues. Almost all kinases contain a conserved 250-300 amino acid catalytic domain containing specific residues and sequence motifs characteristic of the kinase family. The protein kinase catalytic domain can be further divided into 11 subdomains. N-terminal subdomains I-IV fold into a two-lobed structure which binds and orients the ATP donor molecule, and subdomain V spans the two lobes. C-terminal subdomains VI-XI bind the protein substrate and transfer the gamma phosphate from ATP to the hydroxyl group of a tyrosine, serine, or threonine residue. Each of the 11 subdomains contains specific catalytic residues or amino acid motifs characteristic of that subdomain. For example, subdomain I contains an 8-amino acid glycine-rich ATP binding consensus motif, subdomain II contains a critical lysine residue required for maximal catalytic activity, and subdomains VI through IX comprise the highly conserved catalytic core. PTKs and STKs also contain distinct sequence motifs in subdomains VI and VIII which may confer hydroxyamino acid specificity.

In addition, kinases may also be classified by additional amino acid sequences, generally between 5 and 100 residues, which either flank or occur within the kinase domain. These additional amino acid sequences regulate kinase activity and determine substrate specificity. (Reviewed in Hardie, G. and S. Hanks (1995) The Protein Kinase Facts Book, Vol I, pp. 17-20 Academic Press, San Diego CA.). In particular, two protein kinase signature sequences have been identified in the kinase domain, the first containing an active site lysine residue involved in ATP binding, and the second containing an aspartate residue important for catalytic activity. If a protein analyzed includes the two protein kinase signatures, the probability of that protein being a protein kinase is close to 100% (PROSITE: PDOC00100, November 1995).

Protein Tyrosine Kinases

Protein tyrosine kinases (PTKs) may be classified as either transmembrane, receptor PTKs or nontransmembrane, nonreceptor PTK proteins. Transmembrane tyrosine kinases function as receptors for most growth factors. Growth factors bind to the receptor tyrosine kinase (RTK), which causes the receptor to phosphorylate itself (autophosphorylation) and specific intracellular second messenger proteins. Growth factors (GF) that associate with receptor PTKs include epidermal GF, platelet-derived GF, fibroblast GF, hepatocyte GF, insulin and insulin-like GFs, nerve GF, vascular endothelial GF, and macrophage colony stimulating factor.

Nontransmembrane, nonreceptor PTKs lack transmembrane regions and, instead, form signaling complexes with the cytosolic domains of plasma membrane receptors. Receptors that function through non-receptor PTKs include those for cytokines and hormones (growth hormone and prolactin), and antigen-specific receptors on T and B lymphocytes.

Many PTKs were first identified as oncogene products in cancer cells in which PTK activation was no longer subject to normal cellular controls. In fact, about one third of the known oncogenes encode PTKs. Furthermore, cellular transformation (oncogenesis) is often accompanied by increased tyrosine phosphorylation activity (Charbonneau, H. and N.K. Tonks (1992) Annu. Rev. Cell Biol. 8:463-493). Regulation of PTK activity may therefore be an important strategy in controlling some types of cancer.

Protein Serine/Threonine Kinases

Protein serine/threonine kinases (STKs) are nontransmembrane proteins. A subclass of STKs are known as ERKs (extracellular signal regulated kinases) or MAPs (mitogen-activated protein kinases) and are activated after cell stimulation by a variety of hormones and growth factors. Cell stimulation induces a signaling cascade leading to phosphorylation of MEK (MAP/ERK kinase) which, in turn, activates ERK via serine and threonine phosphorylation. A varied number of proteins represent the downstream effectors for the active ERK and implicate it in the control of cell proliferation and differentiation, as well as regulation of the cytoskeleton. Activation of ERK is normally transient, and cells possess dual specificity phosphatases that are responsible for its down-regulation. Also, numerous studies have shown that elevated ERK activity is associated with some cancers. Other STKs include the second messenger dependent protein kinases such as the cyclic-AMP dependent protein kinases (PKA), calcium-calmodulin (CaM) dependent protein kinases, and the mitogen-activated protein kinases (MAP); the cyclin-dependent protein kinases; checkpoint and cell cycle kinases; Numb-associated kinase (Nak); human Fused (hFu); proliferation-related kinases; 5'-AMP-activated protein kinases; and kinases involved in apoptosis.

One member of the ERK family of MAP kinases, ERK 7, is a novel 61-kDa protein that has motif similarities to ERK1 and ERK2, but is not activated by extracellular stimuli as are ERK1 and ERK2 nor by the common activators, c-Jun N-terminal kinase (JNK) and p38 kinase. ERK7 regulates its nuclear localization and inhibition of growth through its C-terminal tail, not through the kinase domain as is typical with other MAP kinases (Abe, M.K. (1999) Mol. Cell. Biol. 19:1301-1312).

The second messenger dependent protein kinases primarily mediate the effects of second messengers such as cyclic AMP (cAMP), cyclic GMP, inositol triphosphate, phosphatidylinositol, 3,4,5-triphosphate, cyclic ADP ribose, arachidonic acid, diacylglycerol and calcium-calmodulin. The PKAs are involved in mediating hormone-induced cellular responses and are activated by cAMP produced within the cell in response to hormone stimulation. cAMP is an intracellular mediator of hormone action in all animal cells that have been studied. Hormone-induced cellular responses include thyroid hormone secretion, cortisol secretion, progesterone secretion, glycogen breakdown, bone resorption, and regulation of heart rate and force of heart muscle contraction. PKA is found in all animal cells and is thought to account for the effects of cAMP in most of these cells. Altered PKA expression is implicated in a variety of disorders and diseases including cancer, thyroid disorders, diabetes, atherosclerosis, and cardiovascular disease (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York NY, pp. 416-431, 1887).

The casein kinase I (CKI) gene family is another subfamily of serine/threonine protein kinases. This continuously expanding group of kinases have been implicated in the regulation of numerous cytoplasmic and nuclear processes, including cell metabolism and DNA replication and repair. CKI enzymes are present in the membranes, nucleus, cytoplasm and cytoskeleton of eukaryotic cells, and on the mitotic spindles of mammalian cells (Fish, K.J. et al. (1995) J. Biol. Chem. 270:14875-14883).

The CKI family members all have a short amino-terminal domain of 9-76 amino acids, a highly conserved kinase domain of 284 amino acids, and a variable carboxyl-terminal domain that ranges from 24 to over 200 amino acids in length (Cegielska, A. et al. (1998) J. Biol. Chem. 273:1357-1364). The CKI family is comprised of highly related proteins, as seen by the identification of isoforms of casein kinase I from a variety of sources. There are at least five mammalian isoforms, α, β, γ, δ, and ε. Fish et al. identified CKI-epsilon from a human placenta cDNA library. It is a basic protein of 416 amino acids and is closest to CKI-delta. Through recombinant expression, it was determined to phosphorylate known CKI substrates and was inhibited by the CKI-specific inhibitor CKI-7. The human gene for CKI-epsilon was able to rescue yeast with a slow-growth phenotype caused by deletion of the yeast CKI locus, HRR250 (Fish et al., *supra*).

The mammalian circadian mutation tau was found to be a semidominant autosomal allele of CKI-epsilon that markedly shortens period length of circadian rhythms in Syrian hamsters. The tau locus is encoded by casein kinase I-epsilon, which is also a homolog of the *Drosophila* circadian gene double-time. Studies of both the wildtype and tau mutant CKI-epsilon enzyme indicated that the mutant enzyme has a noticeable reduction in the maximum velocity and autophosphorylation state. Further, *in vitro,* CKI-epsilon is able to interact with mammalian PERIOD proteins, while the mutant enzyme is deficient in its ability to phosphorylate PERIOD. Lowrey et al. have proposed that CKI-epsilon plays a major role in delaying the negative feedback signal within the transcription-translation-based autoregulatory loop that composes the core of the circadian mechanism. Therefore the CKI-epsilon enzyme is an ideal target for pharmaceutical compounds influencing circadian rhythms, jet-lag and sleep, in addition to other physiologic and metabolic processes under circadian regulation (Lowrey, P.L. et al. (2000) Science 288:483-491).

Homeodomain-interacting protein kinases (HIPKs) are serine/threonine kinases and novel members of the DYRK kinase subfamily (Hofmann, T.G. et al. (2000) Biochimie 82:1123-1127). HIPKs contain a conserved protein kinase domain separated from a domain that interacts with homeoproteins. HIPKs are nuclear kinases, and HIPK2 is highly expressed in neuronal tissue (Kim, Y.H. et al. (1998) J. Biol. Chem. 273:25875-25879; Wang, Y. et al. (2001) Biochim. Biophys. Acta 1518:168-172). HIPKs act as corepressors for homeodomian transcription factors. This corepressor activity is seen in posttranslational modifications such as ubiquitination and phosphorylation, each of which are important in the regulation of cellular protein function (Kim, Y.H. et al. (1999) Proc. Natl. Acad. Sci. USA 96:12350-12355).

The human h-warts protein, a homolog of *Drosophila* warts tumor suppressor gene, maps to chromosome 6q24-25.1. It has a serine/threonine kinase domain and is localized to centrosomes in interphase cells. It is involved in mitosis and functions as a component of the mitotic apparatus (Nishiyama, Y. et al. (1999) FEBS Lett. 459:159-165).

Calcium-Calmodulin Dependent Protein Kinases

Calcium-calmodulin dependent (CaM) kinases are involved in regulation of smooth muscle contraction, glycogen breakdown (phosphorylase kinase), and neurotransmission (CaM kinase I and CaM kinase II). CaM dependent protein kinases are activated by calmodulin, an intracellular calcium receptor, in response to the concentration of free calcium in the cell. Many CaM kinases are also activated by phosphorylation. Some CaM kinases are also activated by autophosphorylation or by other regulatory kinases. CaM kinase I phosphorylates a variety of substrates including the neurotransmitter-related proteins synapsin I and II, the gene transcription regulator, CREB, and the cystic fibrosis conductance regulator protein, CFTR (Haribabu, B. et al. (1995) EMBO J. 14:3679-3686). CaM kinase II also phosphorylates synapsin at different sites and controls the synthesis of catecholamines in the brain through phosphorylation and activation of tyrosine hydroxylase. CaM kinase II controls the synthesis of catecholamines and seratonin, through phosphorylation/activation of tyrosine hydroxylase and tryptophan hydroxylase, respectively (Fujisawa, H. (1990) BioEssays 12:27-29). The mRNA encoding a calmodulin-binding protein kinase-like protein was found to be enriched in mammalian forebrain. This protein is associated with vesicles in both axons and dendrites and accumulates largely postnatally. The amino acid sequence of this protein is similar to CaM-dependent STKs, and the protein binds calmodulin in the presence of calcium (Godbout, M. et al. (1994) J. Neurosci. 14:1-13).

Mitogen-Activated Protein Kinases

The mitogen-activated protein kinases (MAP), which mediate signal transduction from the cell surface to the nucleus via phosphorylation cascades, are another STK family that regulates intracellular signaling pathways. Several subgroups have been identified, and each manifests different substrate specificities and responds to distinct extracellular stimuli (Egan, S.E. and R.A. Weinberg (1993) Nature 365:781-783). There are three kinase modules comprising the MAP kinase cascade: MAPK (MAP), MAPK kinase (MAP2K, MAPKK, or MKK), and MKK kinase (MAP3K, MAPKKK, OR MEKK) (Wang,X.S. et al (1998) Biochem. Biophys. Res. Commun. 253:33-37). The extracellular-regulated kinase (ERK) pathway is activated by growth factors and mitogens, for example, epidermal growth factor (EGF), ultraviolet light, hyperosmolar medium, heat shock, or endotoxic lipopolysaccharide (LPS). The closely related though distinct parallel pathways, the c-Jun N-terminal kinase (JNK), or stress-activated kinase (SAPK) pathway, and the p38 kinase pathway are activated by stress stimuli and proinflammatory cytokines such as tumor necrosis factor (TNF) and interleukin-1 (IL-1). Altered MAP kinase expression is implicated in a variety of disease conditions including cancer, inflammation, immune disorders, and disorders affecting growth and development.. MAP kinase signaling pathways are present in mammalian cells as well as in yeast.

Cyclin-Dependent Protein Kinases

The cyclin-dependent protein kinases (CDKs) are STKs that control the progression of cells through the cell cycle. The entry and exit of a cell from mitosis are regulated by the synthesis and destruction of a family of activating proteins called cyclins. Cyclins are small regulatory proteins that bind to and activate CDKs, which then phosphorylate and activate selected proteins involved in the mitotic process. CDKs are unique in that they require multiple inputs to become activated. In addition to cyclin binding, CDK activation requires the phosphorylation of a specific threonine residue and the dephosphorylation of a specific tyrosine residue on the CDK.

Another family of STKs associated with the cell cycle are the NIMA (never in mitosis)-related kinases (Neks). Both CDKs and Neks are involved in duplication, maturation, and separation of the microtubule organizing center, the centrosome, in animal cells (Fry, A.M. et al. (1998) EMBO J. 17:470-481).

Checkpoint and Cell Cycle Kinases

In the process of cell division, the order and timing of cell cycle transitions are under control of cell cycle checkpoints, which ensure that critical events such as DNA replication and chromosome segregation are carried out with precision. If DNA is damaged, e.g. by radiation, a checkpoint pathway is activated that arrests the cell cycle to provide time for repair. If the damage is extensive, apoptosis is induced. In the absence of such checkpoints, the damaged DNA is inherited by aberrant cells which may cause proliferative disorders such as cancer. Protein kinases play an important role in this process. For example, a specific kinase, checkpoint kinase 1 (Chk1), has been identified in yeast and mammals, and is activated by DNA damage in yeast. Activation of Chk1 leads to the arrest of the cell at the G2/M transition (Sanchez, Y. et al. (1997) Science 277:1497-1501). Specifically, Chk1 phosphorylates the cell division cycle phosphatase CDC25, inhibiting its normal function which is to dephosphorylate and activate the cyclin-dependent kinase Cdc2. Cdc2 activation controls the entry of cells into mitosis (Peng, C.-Y. et al. (1997) Science 277:1501-1505). Thus, activation of Chk1 prevents the damaged cell from entering mitosis. A deficiency in a checkpoint kinase, such as Chk1, may also contribute to cancer by failure to arrest cells with damaged DNA at other checkpoints such as G2/M.

Proliferation-Related Kinases

Proliferation-related kinase is a serum/cytokine inducible STK that is involved in regulation of the cell cycle and cell proliferation in human megakarocytic cells (Li, B. et al. (1996) J. Biol. Chem. 271:19402-19408). Proliferation-related kinase is related to the polo (derived from *Drosophila* polo gene) family of STKs implicated in cell division. Proliferation-related kinase is downregulated in lung tumor tissue and may be a proto-oncogene whose deregulated expression in normal tissue leads to oncogenic transformation.

5'-AMP-activated protein kinase

A ligand-activated STK protein kinase is 5'-AMP-activated protein kinase (AMPK) (Gao, G. et al. (1996) J. Biol Chem. 271:8675-8681). Mammalian AMPK is a regulator of fatty acid and sterol synthesis through phosphorylation of the enzymes acetyl-CoA carboxylase and hydroxymethylglutaryl-CoA reductase and mediates responses of these pathways to cellular stresses such as heat shock and depletion of glucose and ATP. AMPK is a heterotrimeric complex comprised of a catalytic alpha subunit and two non-catalytic beta and gamma subunits that are believed to regulate the activity of the alpha subunit. Subunits of AMPK have a much wider distribution in non-lipogenic tissues such as brain, heart, spleen, and lung than expected. This distribution suggests that its role may extend beyond regulation of lipid metabolism alone.

Kinases in Apoptosis

Apoptosis is a highly regulated signaling pathway leading to cell death that plays a crucial role in tissue development and homeostasis. Deregulation of this process is associated with the pathogenesis of a number of diseases including autoimmune diseases, neurodegenerative disorders, and cancer. Various STKs play key roles in this process. ZIP kinase is an STK containing a C-terminal leucine zipper domain in addition to its N-terminal protein kinase domain. This C-terminal domain appears to mediate homodimerization and activation of the kinase as well as interactions with transcription factors such as activating transcription factor, ATF4, a member of the cyclic-AMP responsive element binding protein (ATF/CREB) family of transcriptional factors (Sanjo, H. et al. (1998) J. Biol. Chem. 273:29066-29071). DRAK1 and DRAK2 are STKs that share homology with the death-associated protein kinases (DAP kinases), known to function in interferon-γ induced apoptosis (Sanjo et al., *supra*). Like ZIP kinase, DAP kinases contain a C-terminal protein-protein interaction domain, in the form of ankyrin repeats, in addition to the N-terminal kinase domain. ZIP, DAP, and DRAK kinases induce morphological changes associated with apoptosis when transfected into NIH3T3 cells (Sanjo et al., *supra*). However, deletion of either the N-terminal kinase catalytic domain or the C-terminal domain of these proteins abolishes apoptosis activity, indicating that in addition to the kinase activity, activity in the C-terminal domain is also necessary for apoptosis, possibly as an interacting domain with a regulator or a specific substrate.

RICK is another STK recently identified as mediating a specific apoptotic pathway involving the death receptor, CD95 (Inohara, N. et al. (1998) J. Biol. Chem. 273:12296-12300). CD95 is a member of the tumor necrosis factor receptor superfamily and plays a critical role in the regulation and homeostasis of the immune system (Nagata, S. (1997) Cell 88:355-365). The CD95 receptor signaling pathway involves recruitment of various intracellular molecules to a receptor complex following ligand binding. This process includes recruitment of the cysteine protease caspase-8 which, in turn, activates a caspase cascade leading to cell death. RICK is composed of an N-terminal kinase catalytic domain and a C-terminal "caspase-recruitment" domain that interacts with caspase-like domains, indicating that RICK plays a role in the recruitment of caspase-8. This interpretation is supported by the fact that the expression of RICK in human 293T cells promotes activation of caspase-8 and potentiates the induction of apoptosis by various proteins involved in the CD95 apoptosis pathway (Inohara et al., *supra*).

Mitochondrial Protein Kinases

A novel class of eukaryotic kinases, related by sequence to prokaryotic histidine protein kinases, are the mitochondrial protein kinases (MPKs) which seem to have no sequence similarity with other eukaryotic protein kinases. These protein kinases are located exclusively in the mitochondrial matrix space and may have evolved from genes originally present in respiration-dependent bacteria which were endocytosed by primitive eukaryotic cells. MPKs are responsible for phosphorylation and inactivation of the branched-chain alpha-ketoacid dehydrogenase and pyruvate dehydrogenase complexes (Harris, R.A. et al. (1995) Adv. Enzyme Regul. 34:147-162). Five MPKs have been identified. Four members correspond to pyruvate dehydrogenase kinase isozymes, regulating the activity of the pyruvate dehydrogenase complex, which is an important regulatory enzyme at the interface between glycolysis and the citric acid cycle. The fifth member corresponds to a branched-chain alpha-ketoacid dehydrogenase kinase, important in the regulation of the pathway for the disposal of branched-chain amino acids. (Harris, R.A. et al. (1997) Adv. Enzyme Regul. 37:271-293). Both starvation and the diabetic state are known to result in a great increase in the activity of the pyruvate dehydrogenase kinase in the liver, heart and muscle of the rat. This increase contributes in both disease states to the phosphorylation and inactivation of the pyruvate dehydrogenase complex and conservation of pyruvate and lactate for gluconeogenesis (Harris (1995) *supra*).

KINASES WITH NON-PROTEIN SUBSTRATES

Lipid and Inositol kinases

Lipid kinases phosphorylate hydroxyl residues on lipid head groups. A family of kinases involved in phosphorylation of phosphatidylinositol (PI) has been described, each member phosphorylating a specific carbon on the inositol ring (Leevers, S.J. et al. (1999) Curr. Opin. Cell. Biol. 11:219-225). The phosphorylation of phosphatidylinositol is involved in activation of the protein kinase C signaling pathway. The inositol phospholipids (phosphoinositides) intracellular signaling pathway begins with binding of a signaling molecule to a G-protein linked receptor in the plasma membrane. This leads to the phosphorylation of phosphatidylinositol (PI) residues on the inner side of the plasma membrane by inositol kinases, thus converting PI residues to the biphosphate state (PIP_2). PIP_2 is then cleaved into inositol triphosphate (IP_3) and diacylglycerol. These two products act as mediators for separate signaling pathways. Cellular responses that are mediated by these pathways are glycogen breakdown in the liver in response to vasopressin, smooth muscle contraction in response to acetylcholine, and thrombin-induced platelet aggregation.

PI 3-kinase (PI3K), which phosphorylates the D3 position of PI and its derivatives, has a central role in growth factor signal cascades involved in cell growth, differentiation, and metabolism. PI3K is a heterodimer consisting of an adapter subunit and a catalytic subunit. The adapter subunit acts as a scaffolding protein, interacting with specific tyrosine-phosphorylated proteins, lipid moieties, and other cytosolic factors. When the adapter subunit binds tyrosine phosphorylated targets, such as the insulin responsive substrate (IRS)-1, the catalytic subunit is activated and converts PI (4,5) bisphosphate (PIP_2) to PI (3,4,5) P_3 (PIP_3). PIP_3 then activates a number of other proteins, including PKA, protein kinase B (PKB), protein kinase C (PKC), glycogen synthase kinase (GSK)-3, and p70 ribosomal s6 kinase. PI3K also interacts directly with the cytoskeletal organizing proteins, Rac, rho, and cdc42 (Shepherd, P.R. et al. (1998) Biochem. J. 333:471-490). Animal models for diabetes, such as *obese* and *fat* mice, have altered PI3K adapter subunit levels. Specific mutations in the adapter subunit have also been found in an insulin-resistant Danish population, suggesting a role for PI3K in type-2 diabetes (Shepard, *supra*).

An example of lipid kinase phosphorylation activity is the phosphorylation of D-erythro-sphingosine to the sphingolipid metabolite, sphingosine-1-phosphate (SPP). SPP has emerged as a novel lipid second-messenger with both extracellular and intracellular actions (Kohama, T. et al. (1998) J. Biol. Chem. 273:23722-23728). Extracellularly, SPP is a ligand for the G-protein coupled receptor EDG-1 (endothelial-derived, G-protein coupled receptor). Intracellularly, SPP regulates cell growth, survival, motility, and cytoskeletal changes. SPP levels are regulated by sphingosine kinases that specifically phosphorylate D-erythro-sphingosine to SPP. The importance of sphingosine kinase in cell signaling is indicated by the fact that various stimuli, including platelet-derived growth factor (PDGF), nerve growth factor, and activation of protein kinase C, increase cellular levels of SPP by activation of sphingosine kinase, and the fact that competitive inhibitors of the enzyme selectively inhibit cell proliferation induced by PDGF (Kohama et al., *supra*).

Purine Nucleotide Kinases

The purine nucleotide kinases, adenylate kinase (ATP:AMP phosphotransferase, or AdK) and guanylate kinase (ATP:GMP phosphotransferase, or GuK) play a key role in nucleotide metabolism and are crucial to the synthesis and regulation of cellular levels of ATP and GTP, respectively. These two molecules are precursors in DNA and RNA synthesis in growing cells and provide the primary source of biochemical energy in cells (ATP), and signal transduction pathways (GTP). Inhibition of various steps in the synthesis of these two molecules has been the basis of many antiproliferative drugs for cancer and antiviral therapy (Pillwein, K. et al. (1990) Cancer Res. 50:1576-1579).

AdK is found in almost all cell types and is especially abundant in cells having high rates of ATP synthesis and utilization such as skeletal muscle. In these cells AdK is physically associated with mitochondria and myofibrils, the subcellular structures that are involved in energy production and utilization, respectively. Recent studies have demonstrated a major function for AdK in transferring high energy phosphoryls from metabolic processes generating ATP to cellular components consuming ATP (Zeleznikar, R.J. et al. (1995) J. Biol. Chem. 270:7311-7319). Thus AdK may have a pivotal role in maintaining energy production in cells, particularly those having a high rate of growth or metabolism such as cancer cells, and may provide a target for suppression of its activity in order to treat certain cancers. Alternatively, reduced AdK activity may be a source of various metabolic, muscle-energy disorders that can result in cardiac or respiratory failure and may be treatable by increasing AdK activity.

GuK, in addition to providing a key step in the synthesis of GTP for RNA and DNA synthesis, also fulfills an essential function in signal transduction pathways of cells through the regulation of GDP and GTP. Specifically, GTP binding to membrane associated G proteins mediates the activation of cell receptors, subsequent intracellular activation of adenyl cyclase, and production of the second messenger, cyclic AMP. GDP binding to G proteins inhibits these processes. GDP and GTP levels also control the activity of certain oncogenic proteins such as $p21^{ras}$ known to be involved in control of cell proliferation and oncogenesis (Bos, J.L. (1989) Cancer Res. 49:4682-4689). High ratios of GTP:GDP caused by suppression of GuK cause activation of $p21^{ras}$ and promote oncogenesis. Increasing GuK activity to increase levels of GDP and reduce the GTP:GDP ratio may provide a therapeutic strategy to reverse oncogenesis.

GuK is an important enzyme in the phosphorylation and activation of certain antiviral drugs useful in the treatment of herpes virus infections. These drugs include the guanine homologs acyclovir and buciclovir (Miller, W.H. and R.L. Miller (1980) J. Biol. Chem. 255:7204-7207; Stenberg, K. et al. (1986) J. Biol. Chem. 261:2134-2139). Increasing GuK activity in infected cells may provide a therapeutic strategy for augmenting the effectiveness of these drugs and possibly for reducing the necessary dosages of the drugs.

Pyrimidine Kinases

The pyrimidine kinases are deoxycytidine kinase and thymidine kinase 1 and 2. Deoxycytidine kinase is located in the nucleus, and thymidine kinase 1 and 2 are found in the cytosol (Johansson, M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11941-11945). Phosphorylation of deoxyribonucleosides by pyrimidine kinases provides an alternative pathway for *de novo* synthesis of DNA precursors. The role of pyrimidine kinases, like purine kinases, in phosphorylation is critical to the activation of several chemotherapeutically important nucleoside analogues (Arner E.S. and S. Eriksson (1995) Pharmacol. Ther. 67:155-186).

PHOSPHATASES

Protein phosphatases are generally characterized as either serine/threonine- or tyrosine-specific based on their preferred phospho-amino acid substrate. However, some phosphatases (DSPs, for dual specificity phosphatases) can act on phosphorylated tyrosine, serine, or threonine residues. The protein serine/threonine phosphatases (PSPs) are important regulators of many cAMP-mediated hormone responses in cells. Protein tyrosine phosphatases (PTPs) play a significant role in cell cycle and cell signaling processes. Another family of phosphatases is the acid phosphatase or histidine acid phosphatase (HAP) family whose members hydrolyze phosphate esters at acidic pH conditions.

PSPs are found in the cytosol, nucleus, and mitochondria and in association with cytoskeletal and membranous structures in most tissues, especially the brain. Some PSPs require divalent cations, such as Ca^{2+} or Mn^{2+}, for activity. PSPs play important roles in glycogen metabolism, muscle contraction, protein synthesis, T cell function, neuronal activity, oocyte maturation, and hepatic metabolism (reviewed in Cohen, P. (1989) Annu. Rev. Biochem. 58:453-508). PSPs can be separated into two classes. The PPP class includes PP1, PP2A, PP2B/calcineurin, PP4, PP5, PP6, and PP7. Members of this class are composed of a homologous catalytic subunit bearing a very highly conserved signature sequence, coupled with one or more regulatory subunits (PROSITE PDOC00115). Further interactions with scaffold and anchoring molecules determine the intracellular localization of PSPs and substrate specificity. The PPM class consists of several closely related isoforms of PP2C and is evolutionarily unrelated to the PPP class.

PP1 dephosphorylates many of the proteins phosphorylated by cyclic AMP-dependent protein kinase (PKA) and is an important regulator of many cAMP-mediated hormone responses in cells. A number of isoforms have been identified, with the alpha and beta forms being produced by alternative splicing of the same gene. Both ubiquitous and tissue-specific targeting proteins for PP1 have been identified. In the brain, inhibition of PP1 activity by the dopamine and adenosine 3',5'-monophosphate-regulated phosphoprotein of 32kDa (DARPP-32) is necessary for normal dopamine response in neostriatal neurons (reviewed in Price, N.E. and M.C. Mumby (1999) Curr. Opin. Neurobiol. 9:336-342). PP1, along with PP2A, has been shown to limit motility in microvascular endothelial cells, suggesting a role for PSPs in the inhibition of angiogenesis (Gabel, S. et al. (1999) Otolaryngol. Head Neck Surg.121:463-468).

PP2A is the main serine/threonine phosphatase. The core PP2A enzyme consists of a single 36 kDa catalytic subunit (C) associated with a 65 kDa scaffold subunit (A), whose role is to recruit additional regulatory subunits (B). Three gene families encoding B subunits are known (PR55, PR61, and PR72), each of which contain multiple isoforms, and additional families may exist (Millward, T.A et al. (1999) Trends Biosci. 24:186-191). These "B-type" subunits are cell type- and tissue-specific and determine the substrate specificity, enzymatic activity, and subcellular localization of the holoenzyme. The PR55 family is highly conserved and bears a conserved motif (PROSITE PDOC00785). PR55 increases PP2A activity toward mitogen-activated protein kinase (MAPK) and MAPK kinase (MEK). PP2A dephosphorylates the MAPK active site, inhibiting the cell's entry into mitosis. Several proteins can compete with PR55 for PP2A core enzyme binding, including the CKII kinase catalytic subunit, polyomavirus middle and small T antigens, and SV40 small t antigen. Viruses may use this mechanism to commandeer PP2A and stimulate progression of the cell through the cell cycle (Pallas, D.C. et al. (1992) J. Virol. 66:886-893). Altered MAP kinase expression is also implicated in a variety of disease conditions including cancer, inflammation, immune disorders, and disorders affecting growth and development. PP2A, in fact, can dephosphorylate and modulate the activities of more than 30 protein kinases *in vitro*, and other evidence suggests that the same is true *in vivo* for such kinases as PKB, PKC, the calmodulin-dependent kinases, ERK family MAP kinases, cyclin-dependent kinases, and the IκB kinases (reviewed in Millward et al., *supra*). PP2A is itself a substrate for CKI and CKII kinases, and can be stimulated by polycationic macromolecules. A PP2A-like phosphatase is necessary to maintain the G1 phase destruction of mammalian cyclins A and B (Bastians, H. et al. (1999) Mol. Biol. Cell 10:3927-3941). PP2A is a major activity in the brain and is implicated in regulating neurofilament stability and normal neural function, particularly the phosphorylation of the microtubule-associated protein tau. Hyperphosphorylation of tau has been proposed to lead to the neuronal degeneration seen in Alzheimer's disease (reviewed in Price and Mumby, *supra*).

PP2B, or calcineurin, is a Ca^{2+}-activated dimeric phosphatase and is particularly abundant in the brain. It consists of catalytic and regulatory subunits, and is activated by the binding of the calcium/calmodulin complex. Calcineurin is the target of the immunosuppressant drugs cyclosporine and FK506. Along with other cellular factors, these drugs interact with calcineurin and inhibit phosphatase activity. In T cells, this blocks the calcium dependent activation of the NF-AT family of transcription factors, leading to immunosuppression. This family is widely distributed, and it is likely that calcineurin regulates gene expression in other tissues as well. In neurons, calcineurin modulates functions which range from the inhibition of neurotransmitter release to desensitization of postsynaptic NMDA-receptor coupled calcium channels to long term memory (reviewed in Price and Mumby, *supra*).

Other members of the PPP class have recently been identified (Cohen, P.T. (1997) Trends Biochem. Sci. 22:245-251). One of them, PP5, contains regulatory domains with tetratricopeptide repeats. It can be activated by polyunsaturated fatty acids and anionic phospholipids *in vitro* and appears to be involved in a number of signaling pathways, including those controlled by atrial natriuretic peptide or steroid hormones (reviewed in Andreeva, A.V. and M.A. Kutuzov (1999) Cell Signal. 11:555-562).

PP2C is a ~42kDa monomer with broad substrate specificity and is dependent on divalent cations (mainly Mn^{2+} or Mg^{2+}) for its activity. PP2C proteins share a conserved N-terminal region with an invariant DGH motif, which contains an aspartate residue involved in cation binding (PROSITE PDOC00792). Targeting proteins and mechanisms regulating PP2C activity have not been identified. PP2C has been shown to inhibit the stress-responsive p38 and Jun kinase (JNK) pathways (Takekawa, M. et al. (1998) EMBO J. 17:4744-4752).

In contrast to PSPs, tyrosine-specific phosphatases (PTPs) are generally monomeric proteins of very diverse size (from 20kDa to greater than 100kDa) and structure that function primarily in the transduction of signals across the plasma membrane. PTPs are categorized as either soluble phosphatases or transmembrane receptor proteins that contain a phosphatase domain. All PTPs share a conserved catalytic domain of about 300 amino acids which contains the active site. The active site consensus sequence includes a cysteine residue which executes a nucleophilic attack on the phosphate moiety during catalysis (Neel, B.G. and N.K. Tonks (1997) Curr. Opin. Cell Biol. 9:193-204). Receptor PTPs are made up of an N-terminal extracellular domain of variable length, a transmembrane region, and a cytoplasmic region that generally contains two copies of the catalytic domain. Although only the first copy seems to have enzymatic activity, the second copy apparently affects the substrate specificity of the first. The extracellular domains of some receptor PTPs contain fibronectin-like repeats, immunoglobulin-like domains, MAM domains (an extracellular motif likely to have an adhesive function), or carbonic anhydrase-like domains (PROSITE PDOC 00323). This wide variety of structural motifs accounts for the diversity in size and specificity of PTPs.

PTPs play important roles in biological processes such as cell adhesion, lymphocyte activation, and cell proliferation. PTPs μ and κ are involved in cell-cell contacts, perhaps regulating cadherin/catenin function. A number of PTPs affect cell spreading, focal adhesions, and cell motility, most of them via the integrin/tyrosine kinase signaling pathway (reviewed in Neel and Tonks, *supra*). CD45 phosphatases regulate signal transduction and lymphocyte activation (Ledbetter, J.A. et al. (1988) Proc. Natl. Acad. Sci. USA 85:8628-8632). Soluble PTPs containing Src-homology-2 domains have been identified (SHPs), suggesting that these molecules might interact with receptor tyrosine kinases. SHP-1 regulates cytokine receptor signaling by controlling the Janus family PTKs in hematopoietic cells, as well as signaling by the T-cell receptor and c-Kit (reviewed in Neel and Tonks, *supra*). M-phase inducer phosphatase plays a key role in the induction of mitosis by dephosphorylating and activating the PTK CDC2, leading to cell division (Sadhu, K. et al. (1990) Proc. Natl. Acad. Sci. USA 87:5139-5143). In addition, the genes encoding at least eight PTPs have been mapped to chromosomal regions that are translocated or rearranged in various neoplastic conditions, including lymphoma, small cell lung carcinoma, leukemia, adenocarcinoma, and neuroblastoma (reviewed in Charbonneau, H. and N.K. Tonks (1992) Annu. Rev. Cell Biol. 8:463-493). The PTP enzyme active site comprises the consensus sequence of the MTM1 gene family. The MTM1 gene is responsible for X-linked recessive myotubular myopathy, a congenital muscle disorder that has been linked to Xq28 (Kioschis, P. et al., (1998) Genomics 54:256-266). Many PTKs are encoded by oncogenes, and it is well known that oncogenesis is often accompanied by increased tyrosine phosphorylation activity. It is therefore possible that PTPs may serve to prevent or reverse cell transformation and the growth of various cancers by controlling the levels of tyrosine phosphorylation in cells. This is supported by studies showing that overexpression of PTP can suppress transformation in cells and that specific inhibition of PTP can enhance cell transformation (Charbonneau and Tonks, *supra*).

Dual specificity phosphatases (DSPs) are structurally more similar to the PTPs than the PSPs. DSPs bear an extended PTP active site motif with an additional 7 amino acid residues. DSPs are primarily associated with cell proliferation and include the cell cycle regulators cdc25A, B, and C. The phosphatases DUSP1 and DUSP2 inactivate the MAPK family members ERK (extracellular signal-regulated kinase), JNK (c-Jun N-terminal kinase), and p38 on both tyrosine and threonine residues (PROSITE PDOC 00323, *supra*). In the activated state, these kinases have been implicated in neuronal differentiation, proliferation, oncogenic transformation, platelet aggregation, and apoptosis. Thus, DSPs are necessary for proper regulation of these processes (Muda, M. et al. (1996) J. Biol. Chem. 271:27205-27208). The tumor suppressor PTEN is a DSP that also shows lipid phosphatase activity. It seems to negatively regulate interactions with the extracellular matrix and maintains sensitivity to apoptosis. PTEN has been implicated in the prevention of angiogenesis (Giri, D. and M. Ittmann (1999) Hum. Pathol. 30:419-424) and abnormalities in its expression are associated with numerous cancers (reviewed in Tamura, M. et al. (1999) J. Natl. Cancer Inst. 91:1820-1828).

Histidine acid phosphatase (HAP; EXPASY EC 3.1.3.2), also known as acid phosphatase, hydrolyzes a wide spectrum of substrates including alkyl, aryl, and acyl orthophosphate monoesters and phosphorylated proteins at low pH. HAPs share two regions of conserved sequences, each centered around a histidine residue which is involved in catalytic activity. Members of the HAP family include lysosomal acid phosphatase (LAP) and prostatic acid phosphatase (PAP), both sensitive to inhibition by L-tartrate (PROSITE PDOC00538).

Synaptojanin, a polyphosphoinositide phosphatase, dephosphorylates phosphoinositides at positions 3, 4 and 5 of the inositol ring. Synaptojanin is a major presynaptic protein found at clathrin-coated endocytic intermediates in nerve terminals, and binds the clathrin coat-associated protein, EPS15. This binding is mediated by the C-terminal region of synaptojanin-170, which has 3 Asp-Pro-Phe amino acid repeats. Further, this 3 residue repeat had been found to be the binding site for the EH domains of EPS15 (Haffner, C. et al. (1997) FEBS Lett. 419:175-180). Additionally, synaptojanin may potentially regulate interactions of endocytic proteins with the plasma membrane, and be involved in synaptic vesicle recycling (Brodin, L. et al. (2000) Curr. Opin. Neurobiol. 10:312-320). Studies in mice with a targeted disruption in the synaptojanin 1 gene (Synj1) were shown to support coat formation of endocytic vesicles more effectively than was seen in wild-type mice, suggesting that Synj1 can act as a negative regulator of membrane-coat protein interactions. These findings provide genetic evidence for a crucial role of phosphoinositide metabolism in synaptic vesicle recycling (Cremona, O. et al. (1999) Cell 99:179-188).

Expression profiling

Microarrays are analytical tools used in bioanalysis. A microarray has a plurality of molecules spatially distributed over, and stably associated with, the surface of a solid support. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as gene sequencing, monitoring gene expression, gene mapping, bacterial identification, drug discovery, and combinatorial chemistry.

One area in particular in which microarrays find use is in gene expression analysis. Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

There is a need in the art for new compositions, including nucleic acids and proteins, for the diagnosis, prevention, and treatment of cardiovascular diseases, immune system disorders, neurological disorders, disorders affecting growth and development, lipid disorders, cell proliferative disorders, and cancers.

SUMMARY OF THE INVENTION

Various embodiments of the invention provide purified polypeptides, kinases and phosphatases, referred to collectively as 'KPP' and individually as 'KPP-1,' 'KPP-2,' 'KPP-3,' 'KPP-4,' 'KPP-5,' 'KPP-6,' 'KPP-7,' 'KPP-8,' 'KPP-9,' 'KPP-10,' 'KPP-11,' 'KPP-12,' 'KPP-13,' 'KPP-14,' 'KPP-15,' 'KPP-16,' 'KPP-17,' 'KPP-18,' 'KPP-19,' 'KPP-20,' 'KPP-21,' 'KPP-22,' 'KPP-23,' 'KPP-24,' 'KPP-25,' 'KPP-26,' 'KPP-27,' 'KPP-28,' 'KPP-29,' 'KPP-30,' 'KPP-31,' 'KPP-32,' 'KPP-33,' 'KPP-34,' 'KPP-35,' 'KPP-36,' 'KPP-37,' 'KPP-38,' 'KPP-39,' 'KPP-40,' 'KPP-41,' 'KPP-42,' and methods for using these proteins and their encoding polynucleotides for the detection, diagnosis, and treatment of diseases and medical conditions. Embodiments also provide methods for utilizing the purified kinases and phosphatases and/or their encoding polynucleotides for facilitating the drug discovery process, including determination of efficacy, dosage, toxicity, and pharmacology. Related embodiments provide methods for utilizing the purified kinases and phosphatases and/or their encoding polynucleotides for investigating the pathogenesis of diseases and medical conditions.

An embodiment provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. Another embodiment provides an isolated polypeptide comprising an amino acid sequence of SEQ ID NO:1-42.

Still another embodiment provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. In another embodiment, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-42. In an alternative embodiment, the polynucleotide is selected from the group consisting of SEQ ID NO:43-84.

Still another embodiment provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. Another embodiment provides a cell transformed with the recombinant polynucleotide. Yet another embodiment provides a transgenic organism comprising the recombinant polynucleotide.

Another embodiment provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Yet another embodiment provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

Still yet another embodiment provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In other embodiments, the polynucleotide can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex. In a related embodiment, the method can include detecting the amount of the hybridization complex. In still other embodiments, the probe can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Still yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof. In a related embodiment, the method can include detecting the amount of the amplified target polynucleotide or fragment thereof.

Another embodiment provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and a pharmaceutically acceptable excipient. In one embodiment, the composition can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. Other embodiments provide a method of treating a disease or condition associated with decreased or abnormal expression of functional KPP, comprising administering to a patient in need of such treatment the composition.

Yet another embodiment provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. Another embodiment provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with decreased expression of functional KPP, comprising administering to a patient in need of such treatment the composition.

Still yet another embodiment provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. Another embodiment provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with overexpression of functional KPP, comprising administering to a patient in need of such treatment the composition.

Another embodiment provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

Yet another embodiment provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

Still yet another embodiment provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

Another embodiment provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide can comprise a fragment of a polynucleotide selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for full length polynucleotide and polypeptide embodiments of the invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog, and the PROTEOME database identification numbers and annotations of PROTEOME database homologs, for polypeptide embodiments of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide embodiments, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide embodiments, along with selected fragments of the polynucleotides.

Table 5 shows representative cDNA libraries for polynucleotide embodiments.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze polynucleotides and polypeptides, along with applicable descriptions, references, and threshold parameters.

Table 8 shows single nucleotide polymorphisms found in polynucleotide sequences of the invention, along with allele frequencies in different human populations.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleic acids, and methods are described, it is understood that embodiments of the invention are not limited to the particular machines, instruments, materials, and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the invention.

As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with various embodiments of the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"KPP" refers to the amino acid sequences of substantially purified KPP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of KPP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of KPP either by directly interacting with KPP or by acting on components of the biological pathway in which KPP participates.

An "allelic variant" is an alternative form of the gene encoding KPP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding KPP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as KPP or a polypeptide with at least one functional characteristic of KPP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding KPP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide encoding KPP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent KPP. Deliberate amino acid substitutions may be made on the basis of one or more similarities in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of KPP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" can refer to an oligopeptide, a peptide, a polypeptide, or a protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid. Amplification may be carried out using polymerase chain reaction (PCR) technologies or other nucleic acid amplification technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of KPP. Antagonists may include proteins such as antibodies, anticalins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of KPP either by directly interacting with KPP or by acting on components of the biological pathway in which KPP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind KPP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH_2), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker (Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13).

The term "intramer" refers to an aptamer which is expressed *in vivo*. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a polynucleotide having a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic KPP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide" and a "composition comprising a given polypeptide" can refer to any composition containing the given polynucleotide or polypeptide. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotides encoding KPP or fragments of KPP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (Accelrys, Burlington MA) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr

Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of KPP or a polynucleotide encoding KPP which can be identical in sequence to, but shorter in length than, the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from about 5 to about 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:43-84 can comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:43-84, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:43-84 can be employed in one or more embodiments of methods of the invention, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:43-84 from related polynucleotides. The precise length of a fragment of SEQ ID NO:43-84 and the region of SEQ ID NO:43-84 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-42 is encoded by a fragment of SEQ ID NO:43-84. A fragment of SEQ ID NO:1-42 can comprise a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-42. For example, a fragment of SEQ ID NO:1-42 can be used as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-42. The precise length of a fragment of SEQ ID NO:1-42 and the region of SEQ ID NO:1-42 to which the fragment corresponds can be determined based on the intended purpose for the fragment using one or more analytical methods described herein or otherwise known in the art.

A "full length" polynucleotide is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, alternatively, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of identical residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of identical residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide. The phrases "percent similarity" and "% similarity," as applied to polypeptide sequences, refer to the percentage of residue matches, including identical residue matches and conservative substitutions, between at least two polypeptide sequences aligned using a standardized algorithm. In contrast, conservative substitutions are not included in the calculation of percent identity between polypeptide sequences.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. and D.W.

Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY, ch. 9).

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acids by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid present in solution and another nucleic acid immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed)..

The words "insertion" and "addition" refer to changes in an amino acid or polynucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of KPP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of KPP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, antibody, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of KPP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of KPP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an KPP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of KPP.

"Probe" refers to nucleic acids encoding KPP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acids. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in, for example, Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY), Ausubel, F.M. et al. (1999; Short Protocols in Molecular Biology, 4th ed., John Wiley & Sons, New York NY), and Innis, M. et al. (1990; PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a nucleic acid that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook and Russell (*supra*). The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA molecule, is composed of the same linear sequence of nucleotides as the reference DNA molecule with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing KPP, nucleic acids encoding KPP, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably at least about 75% free, and most preferably at least about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In another embodiment, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook and Russell (*supra*).

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotides that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity or sequence similarity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity or sequence similarity over a certain defined length of one of the polypeptides.

THE INVENTION

Various embodiments of the invention include new human kinases and phosphatases (KPP), the polynucleotides encoding KPP, and the use of these compositions for the diagnosis, treatment, or prevention of cardiovascular diseases, immune system disorders, neurological disorders, disorders affecting growth and development, lipid disorders, cell proliferative disorders, and cancers.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide embodiments of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown. Column 6 shows the Incyte ID numbers of physical, full length clones corresponding to the polypeptide and polynucleotide sequences of the invention. The full length clones encode polypeptides which have at least 95% sequence identity to the polypeptide sequences shown in column 3.

Table 2 shows sequences with homology to polypeptide embodiments of the invention as identified by BLAST analysis against the GenBank protein (genpept) database and the PROTEOME database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog and the PROTEOME database identification numbers (PROTEOME ID NO:) of the nearest PROTEOME database homologs. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank and PROTEOME database homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Accelrys, Burlington MA). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are kinases and phosphatases. For example, SEQ ID NO:4 is 100% identical, from residue A26 to residue A279, to human SIRP-beta1 (GenBank ID g2052058) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.0e-136, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:4 also has homology to signal regulatory protein beta-1, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:4 also contains a immunoglobulin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLAST, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:4 is a signal regulatory protein beta-1.

As another example, SEQ ID NO:11 is 98% identical, from residue M1 to residue G341, to human MNK1 (GenBank ID g2077825) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.1E-181, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. As determined by BLAST analysis using the PROTEOME database, SEQ ID NO:11 also has homology to MAP kinase-interacting serine/threonine kinase 1, a kinase which is activated by the ERK and p38 MAP kinase signaling pathways. This kinase (PROTEOME ID 336418|MKNK1) also phosphorylates translation initiation factor 4E (EIF4E), reduces cap-dependent translation, and phosphorylates cytosolic phospholipase. SEQ ID NO:11 also contains a protein kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, PROFILESCAN and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:11 is a MAP kinase-activated protein kinase.

As another example, SEQ ID NO:12 is 99% identical, from residue M1 to residue R498, to human clk2 kinase (GenBank ID g2564911) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.6e-277, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also has homology to proteins that have kinase activity and are members of the LAMMER family of dual specificity protein kinases, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:12 also contains a protein kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, BLAST, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:12 is a member of the kinase family.

As another example, SEQ ID NO:15 is 92% identical, from residue S89 to residue L893, to human homolog of *Drosophila* discs large protein (GenBank ID g558436) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:15 also has homology to proteins that are localized to the plasma membrane, have kinase activity, and are a member of the membrane-associated guanylate kinase (MAGUK) family, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:15 also contains a guanylate kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:15 is a kinase.

As another example, SEQ ID NO:25 is 92% identical, from residue V11 to residue K323, and 99% identical, from residue M362 to residue Q495, to rat calmodulin-dependent protein kinase II gamma subunit (GenBank ID g206152) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.9e-227, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:25 also has homology to proteins that have protein kinase activity, and are calcium/calmodulin-dependent protein kinase II, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:25 also contains a protein kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains and a serine/threonine protein kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based SMRT database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, BLAST against the PRODOM and DOMO databases, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:25 is a protein kinase.

SEQ ID NO:1-3, SEQ ID NO:5-10, SEQ ID NO:13-14, SEQ ID NO:16-24, and SEQ ID

NO:26-42 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-42 are described in Table 7.

As shown in Table 4, the full length polynucleotide embodiments were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide embodiments, and of fragments of the polynucleotides which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:43-84 or that distinguish between SEQ ID NO:43-84 and related polynucleotides.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotides. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_$XXXXXX_N_1_N_2_YYYYY_N_3_N_4$ represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FL*XXXXXX*_*gAAAAA*_*gBBBBB*_1_*N* is a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, *gAAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, *gBBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	**Type of analysis and/or examples of programs**
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotides which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotides. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide sequences of the invention, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO:) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence. Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

The invention also encompasses KPP variants. Various embodiments of KPP variants can have at least about 80%, at least about 90%, or at least about 95% amino acid sequence identity to the KPP amino acid sequence, and can contain at least one functional or structural characteristic of KPP.

Various embodiments also encompass polynucleotides which encode KPP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:43-84, which encodes KPP. The polynucleotide sequences of SEQ ID NO:43-84, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses variants of a polynucleotide encoding KPP. In particular, such a variant polynucleotide will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a polynucleotide encoding KPP. A particular aspect of the invention encompasses a variant of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:43-84 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:43-84. Any one of the polynucleotide variants described above can encode a polypeptide which contains at least one functional or structural characteristic of KPP.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide encoding KPP. A splice variant may have portions which have significant sequence identity to a polynucleotide encoding KPP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to a polynucleotide encoding KPP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide encoding KPP. For example, a polynucleotide comprising a sequence of SEQ ID NO:50 and a polynucleotide comprising a sequence of SEQ ID NO:51 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:62, a polynucleotide comprising a sequence of SEQ ID NO:63 a polynucleotide comprising a sequence of SEQ ID NO:64, and a polynucleotide comprising a sequence of SEQ ID NO:65 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:71 and a polynucleotide comprising a sequence of SEQ ID NO:72 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:74 and a polynucleotide comprising a sequence of SEQ ID NO:82 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:77 and a polynucleotide comprising a sequence of SEQ ID NO:78 are splice variants of each other; and a polynucleotide comprising a sequence of SEQ ID NO:80, a polynucleotide comprising a sequence of SEQ ID NO:81 a polynucleotide comprising a sequence of SEQ ID NO:83, and a polynucleotide comprising a sequence of SEQ ID NO:84 are splice variants of each other. Any one of the splice variants described above can encode a polypeptide which contains at least one functional or structural characteristic of KPP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding KPP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring KPP, and all such variations are to be considered as being specifically disclosed.

Although polynucleotides which encode KPP and its variants are generally capable of hybridizing to polynucleotides encoding naturally occurring KPP under appropriately selected conditions of stringency, it may be advantageous to produce polynucleotides encoding KPP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding KPP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of polynucleotides which encode KPP and KPP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic polynucleotide may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a polynucleotide encoding KPP or any fragment thereof.

Embodiments of the invention can also include polynucleotides that are capable of hybridizing to the claimed polynucleotides, and, in particular, to those having the sequences shown in SEQ ID NO:43-84 and fragments thereof, under various conditions of stringency (Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511). Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Biosciences, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad CA). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Biosciences), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art (Ausubel et al., *supra*, ch. 7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

The nucleic acids encoding KPP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). A third method, capture

PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art (Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotides or fragments thereof which encode KPP may be cloned in recombinant DNA molecules that direct expression of KPP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptides may be produced and used to express KPP.

The polynucleotides of the invention can be engineered using methods generally known in the art in order to alter KPP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of KPP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, polynucleotides encoding KPP may be synthesized, in whole or in part, using one or more chemical methods well known in the art (Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232). Alternatively, KPP itself or a fragment thereof may be synthesized using chemical methods known in the art. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. et al. (1995) Science 269:202-204). Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of KPP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography (Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421). The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing (Creighton, *supra*, pp. 28-53).

In order to express a biologically active KPP, the polynucleotides encoding KPP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotides encoding KPP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of polynucleotides encoding KPP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where a polynucleotide sequence encoding KPP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

Methods which are well known to those skilled in the art may be used to construct expression vectors containing polynucleotides encoding KPP and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination (Sambrook and Russell, *supra*, ch. 1-4, and 8; Ausubel et al., *supra*, ch. 1, 3, and 15).

A variety of expression vector/host systems may be utilized to contain and express polynucleotides encoding KPP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems (Sambrook and Russell, *supra*; Ausubel et al., *supra*; Van

Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355). Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of polynucleotides to the targeted organ, tissue, or cell population (Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5:350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Buller, R.M. et al. (1985) Nature 317:813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31:219-226; Verma, I.M. and N. Somia (1997) Nature 389:239-242). The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotides encoding KPP. For example, routine cloning, subcloning, and propagation of polynucleotides encoding KPP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Invitrogen). Ligation of polynucleotides encoding KPP into the vector's multiple cloning site disrupts the *lac*Z gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509). When large quantities of KPP are needed, e.g. for the production of antibodies, vectors which direct high level expression of KPP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of KPP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign polynucleotide sequences into the host genome for stable propagation (Ausubel et al., *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184).

Plant systems may also be used for expression of KPP. Transcription of polynucleotides encoding KPP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection (The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196).

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, polynucleotides encoding KPP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses KPP in host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355).

For long term production of recombinant proteins in mammalian systems, stable expression of KPP in cell lines is preferred. For example, polynucleotides encoding KPP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively (Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823). Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14). Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051). Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β-glucuronidase and its substrate β-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding KPP is inserted within a marker gene sequence, transformed cells containing polynucleotides encoding KPP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding KPP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the polynucleotide encoding KPP and that express KPP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of KPP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on KPP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art (Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding KPP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, polynucleotides encoding KPP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Biosciences, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with polynucleotides encoding KPP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode KPP may be designed to contain signal sequences which direct secretion of KPP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted polynucleotides or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant polynucleotides encoding KPP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric KPP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of KPP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the KPP encoding sequence and the heterologous protein sequence, so that KPP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In another embodiment, synthesis of radiolabeled KPP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

KPP, fragments of KPP, or variants of KPP may be used to screen for compounds that specifically bind to KPP. One or more test compounds may be screened for specific binding to KPP. In various embodiments, 1, 2, 3, 4, 5, 10, 20, 50, 100, or 200 test compounds can be screened for specific binding to KPP. Examples of test compounds can include antibodies, anticalins, oligonucleotides, proteins (e.g., ligands or receptors), or small molecules.

In related embodiments, variants of KPP can be used to screen for binding of test compounds, such as antibodies, to KPP, a variant of KPP, or a combination of KPP and/or one or more variants KPP. In an embodiment, a variant of KPP can be used to screen for compounds that bind to a variant of KPP, but not to KPP having the exact sequence of a sequence of SEQ ID NO:1-42. KPP variants used to perform such screening can have a range of about 50% to about 99% sequence identity to KPP, with various embodiments having 60%, 70%, 75%, 80%, 85%, 90%, and 95% sequence identity.

In an embodiment, a compound identified in a screen for specific binding to KPP can be closely related to the natural ligand of KPP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner (Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):Chapter 5). In another embodiment, the compound thus identified can be a natural ligand of a receptor KPP (Howard, A.D. et al. (2001) Trends Pharmacol. Sci.22:132-140; Wise, A. et al. (2002) Drug Discovery Today 7:235-246).

In other embodiments, a compound identified in a screen for specific binding to KPP can be closely related to the natural receptor to which KPP binds, at least a fragment of the receptor, or a fragment of the receptor including all or a portion of the ligand binding site or binding pocket. For example, the compound may be a receptor for KPP which is capable of propagating a signal, or a decoy receptor for KPP which is not capable of propagating a signal (Ashkenazi, A. and V.M. Divit (1999) Curr. Opin. Cell Biol. 11:255-260; Mantovani, A. et al. (2001) Trends Immunol. 22:328-336). The compound can be rationally designed using known techniques. Examples of such techniques include those used to construct the compound etanercept (ENBREL; Amgen Inc., Thousand Oaks CA), which is efficacious for treating rheumatoid arthritis in humans. Etanercept is an engineered p75 tumor necrosis factor (TNF) receptor dimer linked to the Fc portion of human IgG_1 (Taylor, P.C. et al. (2001) Curr. Opin. Immunol. 13:611-616).

In one embodiment, two or more antibodies having similar or, alternatively, different specificities can be screened for specific binding to KPP, fragments of KPP, or variants of KPP. The binding specificity of the antibodies thus screened can thereby be selected to identify particular fragments or variants of KPP. In one embodiment, an antibody can be selected such that its binding specificity allows for preferential identification of specific fragments or variants of KPP. In another embodiment, an antibody can be selected such that its binding specificity allows for preferential diagnosis of a specific disease or condition having increased, decreased, or otherwise abnormal production of KPP.

In an embodiment, anticalins can be screened for specific binding to KPP, fragments of KPP, or variants of KPP. Anticalins are ligand-binding proteins that have been constructed based on a lipocalin scaffold (Weiss, G.A. and H.B. Lowman (2000) Chem. Biol. 7:R177-R184; Skerra, A. (2001) J. Biotechnol. 74:257-275). The protein architecture of lipocalins can include a beta-barrel having eight antiparallel beta-strands, which supports four loops at its open end. These loops form the natural ligand-binding site of the lipocalins, a site which can be re-engineered *in vitro* by amino acid substitutions to impart novel binding specificities. The amino acid substitutions can be made using methods known in the art or described herein, and can include conservative substitutions (e.g., substitutions that do not alter binding specificity) or substitutions that modestly, moderately, or significantly alter binding specificity.

In one embodiment, screening for compounds which specifically bind to, stimulate, or inhibit

KPP involves producing appropriate cells which express KPP, either as a secreted protein or on the cell membrane. Preferred cells can include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing KPP or cell membrane fractions which contain KPP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either KPP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with KPP, either in solution or affixed to a solid support, and detecting the binding of KPP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

An assay can be used to assess the ability of a compound to bind to its natural ligand and/or to inhibit the binding of its natural ligand to its natural receptors. Examples of such assays include radio-labeling assays such as those described in U.S. Patent No. 5,914,236 and U.S. Patent No. 6,372,724. In a related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a receptor) to improve or alter its ability to bind to its natural ligands (Matthews, D.J. and J.A. Wells. (1994) Chem. Biol. 1:25-30). In another related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a ligand) to improve or alter its ability to bind to its natural receptors (Cunningham, B.C. and J.A. Wells (1991) Proc. Natl. Acad. Sci. USA 88:3407-3411; Lowman, H.B. et al. (1991) J. Biol. Chem. 266:10982-10988).

KPP, fragments of KPP, or variants of KPP may be used to screen for compounds that modulate the activity of KPP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for KPP activity, wherein KPP is combined with at least one test compound, and the activity of KPP in the presence of a test compound is compared with the activity of KPP in the absence of the test compound. A change in the activity of KPP in the presence of the test compound is indicative of a compound that modulates the activity of KPP. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising KPP under conditions suitable for KPP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of KPP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding KPP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease (see, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337). For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding KPP may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding KPP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding KPP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress KPP, e.g., by secreting KPP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of KPP and kinases and phosphatases. In addition, examples of tissues expressing KPP can be found in Table 6 and can also be found in Example XI. Therefore, KPP appears to play a role in cardiovascular diseases, immune system disorders, neurological disorders, disorders affecting growth and development, lipid disorders, cell proliferative disorders, and cancers. In the treatment of disorders associated with increased KPP expression or activity, it is desirable to decrease the expression or activity of KPP. In the treatment of disorders associated with decreased KPP expression or activity, it is desirable to increase the expression or activity of KPP.

Therefore, in one embodiment, KPP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of KPP. Examples of such disorders include, but are not limited to, a cardiovascular disease such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; an immune system disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a disorder affecting growth and development such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a lipid disorder such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM_2 gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatoses, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, uterus, leukemias such as multiple myeloma, and lymphomas such as Hodgkin's disease.

In another embodiment, a vector capable of expressing KPP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of KPP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified KPP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of KPP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of KPP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of KPP including, but not limited to, those listed above.

In a further embodiment, an antagonist of KPP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of KPP. Examples of such disorders include, but are not limited to, those cardiovascular diseases, immune system disorders, neurological disorders, disorders affecting growth and development, lipid disorders, cell proliferative disorders, and cancers described above. In one aspect, an antibody which specifically binds KPP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express KPP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding KPP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of KPP including, but not limited to, those described above.

In other embodiments, any protein, agonist, antagonist, antibody, complementary sequence, or vector embodiments may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of KPP may be produced using methods which are generally known in the art. In particular, purified KPP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind KPP. Antibodies to KPP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. In an embodiment, neutralizing antibodies (i.e., those which inhibit dimer formation) can be used therapeutically. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have application in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with KPP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to KPP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are substantially identical to a portion of the amino acid sequence of the natural protein. Short stretches of KPP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to KPP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce KPP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for KPP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 246:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between KPP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering KPP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for KPP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of KPP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple KPP epitopes, represents the average affinity, or avidity, of the antibodies for KPP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular KPP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the KPP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of KPP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of KPP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available (Catty, *supra*; Coligan et al., *supra*).

In another embodiment of the invention, polynucleotides encoding KPP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding KPP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding KPP (Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, Totawa NJ).

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein (Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102:469-475; Scanlon, K.J. et al. (1995) 9:1288-1296). Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors (Miller, A.D. (1990) Blood 76:271; Ausubel et al., *supra*; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63:323-347). Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art (Rossi, J.J. (1995) Br. Med. Bull. 51:217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87:1308-1315; Morris, M.C. et al. (1997) Nucleic Acids Res. 25:2730-2736).

In another embodiment of the invention, polynucleotides encoding KPP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in KPP expression or regulation causes disease, the expression of KPP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in KPP are treated by constructing mammalian expression vectors encoding KPP and introducing these vectors by mechanical means into KPP-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of KPP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). KPP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding KPP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to KPP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding KPP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4$^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In an embodiment, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding KPP to cells which have one or more genetic abnormalities with respect to the expression of KPP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999; Annu. Rev. Nutr. 19:511-544) and Verma, I.M. and N. Somia (1997; Nature 18:389:239-242).

In another embodiment, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding KPP to target cells which have one or more genetic abnormalities with respect to the expression of KPP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing KPP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.

169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999; J. Virol. 73:519-532) and Xu, H. et al. (1994; Dev. Biol. 163:152-161). The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another embodiment, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding KPP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for KPP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of KPP-coding RNAs and the synthesis of high levels of KPP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of KPP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of RNA molecules encoding KPP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA molecules encoding KPP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

In other embodiments of the invention, the expression of one or more selected polynucleotides of the present invention can be altered, inhibited, decreased, or silenced using RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) methods known in the art. RNAi is a post-transcriptional mode of gene silencing in which double-stranded RNA (dsRNA) introduced into a targeted cell specifically suppresses the expression of the homologous gene (i.e., the gene bearing the sequence complementary to the dsRNA). This effectively knocks out or substantially reduces the expression of the targeted gene. PTGS can also be accomplished by use of DNA or DNA fragments as well. RNAi methods are described by Fire, A. et al. (1998; Nature 391:806-811) and Gura, T. (2000; Nature 404:804-808). PTGS can also be initiated by introduction of a complementary segment of DNA into the selected tissue using gene delivery and/or viral vector delivery methods described herein or known in the art.

RNAi can be induced in mammalian cells by the use of small interfering RNA also known as siRNA. SiRNA are shorter segments of dsRNA (typically about 21 to 23 nucleotides in length) that result *in vivo* from cleavage of introduced dsRNA by the action of an endogenous ribonuclease. SiRNA appear to be the mediators of the RNAi effect in mammals. The most effective siRNAs appear to be 21 nucleotide dsRNAs with 2 nucleotide 3' overhangs. The use of siRNA for inducing RNAi in mammalian cells is described by Elbashir, S.M. et al. (2001; Nature 411:494-498).

SiRNA can either be generated indirectly by introduction of dsRNA into the targeted cell, or directly by mammalian transfection methods and agents described herein or known in the art (such as liposome-mediated transfection, viral vector methods, or other polynucleotide delivery/introductory methods). Suitable SiRNAs can be selected by examining a transcript of the target polynucleotide (e.g., mRNA) for nucleotide sequences downstream from the AUG start codon and recording the occurrence of each nucleotide and the 3' adjacent 19 to 23 nucleotides as potential siRNA target sites, with sequences having a 21 nucleotide length being preferred. Regions to be avoided for target siRNA sites include the 5' and 3' untranslated regions (UTRs) and regions near the start codon (within 75 bases), as these may be richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP endonuclease complex. The selected target sites for siRNA can then be compared to the appropriate genome database (e.g., human, etc.) using BLAST or other sequence comparison algorithms known in the art. Target sequences with significant homology to other coding sequences can be eliminated from consideration. The selected SiRNAs can be produced by chemical synthesis methods known in the art or by *in vitro* transcription using commercially available methods and kits such as the SILENCER siRNA construction kit (Ambion, Austin TX).

In alternative embodiments, long-term gene silencing and/or RNAi effects can be induced in selected tissue using expression vectors that continuously express siRNA. This can be accomplished using expression vectors that are engineered to express hairpin RNAs (shRNAs) using methods known in the art (see, e.g., Brummelkamp, T.R. et al. (2002) Science 296:550-553; and Paddison, P.J. et al. (2002) Genes Dev. 16:948-958). In these and related embodiments, shRNAs can be delivered to target cells using expression vectors known in the art. An example of a suitable expression vector for delivery of siRNA is the PSILENCER1.0-U6 (circular) plasmid (Ambion). Once delivered to the target tissue, shRNAs are processed *in vivo* into siRNA-like molecules capable of carrying out gene-specific silencing.

In various embodiments, the expression levels of genes targeted by RNAi or PTGS methods can be determined by assays for mRNA and/or protein analysis. Expression levels of the mRNA of a targeted gene, can be determined by northern analysis methods using, for example, the NORTHERNMAX-GLY kit (Ambion); by microarray methods; by PCR methods; by real time PCR methods; and by other RNA/polynucleotide assays known in the art or described herein. Expression levels of the protein encoded by the targeted gene can be determined by Western analysis using standard techniques known in the art.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding KPP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased KPP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding KPP may be therapeutically useful, and in the treatment of disorders associated with decreased KPP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding KPP may be therapeutically useful.

In various embodiments, one or more test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding KPP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding KPP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding KPP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art (Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient.

Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of KPP, antibodies to KPP, and mimetics, agonists, antagonists, or inhibitors of KPP.

In various embodiments, the compositions described herein, such as pharmaceutical compositions, may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery allows administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising KPP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, KPP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example KPP or fragments thereof, antibodies of KPP, and agonists, antagonists or inhibitors of KPP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind KPP may be used for the diagnosis of disorders characterized by expression of KPP, or in assays to monitor patients being treated with KPP or agonists, antagonists, or inhibitors of KPP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for KPP include methods which utilize the antibody and a label to detect KPP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring KPP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of KPP expression. Normal or standard values for KPP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to KPP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of KPP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, polynucleotides encoding KPP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotides, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of KPP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of KPP, and to monitor regulation of KPP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotides, including genomic sequences, encoding KPP or closely related molecules may be used to identify nucleic acid sequences which encode KPP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding KPP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the KPP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:43-84 or from genomic sequences including promoters, enhancers, and introns of the KPP gene.

Means for producing specific hybridization probes for polynucleotides encoding KPP include the cloning of polynucleotides encoding KPP or KPP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotides encoding KPP may be used for the diagnosis of disorders associated with expression of KPP. Examples of such disorders include, but are not limited to, a cardiovascular disease such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; an immune system disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a disorder affecting growth and development such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a lipid disorder such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM_2 gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatoses, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, uterus, leukemias such as multiple myeloma, and lymphomas such as Hodgkin's disease. Polynucleotides encoding KPP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered KPP expression. Such qualitative or quantitative methods are well known in the art.

In a particular embodiment, polynucleotides encoding KPP may be used in assays that detect the presence of associated disorders, particularly those mentioned above. Polynucleotides complementary to sequences encoding KPP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of polynucleotides encoding KPP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of KPP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding KPP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding KPP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding KPP, or a fragment of a polynucleotide complementary to the polynucleotide encoding KPP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from polynucleotides encoding KPP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from polynucleotides encoding KPP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641).

Methods which may also be used to quantify the expression of KPP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves (Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, KPP, fragments of KPP, or antibodies specific for KPP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time (Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484; hereby expressly incorporated by reference herein). Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity

(Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity (see, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm). Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In an embodiment, the toxicity of a test compound can be assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another embodiment relates to the use of the polypeptides disclosed herein to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie

Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of interest. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for KPP to quantify the levels of KPP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art (Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In another embodiment of the invention, nucleic acid sequences encoding KPP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; Trask, B.J. (1991) Trends Genet. 7:149-154). Once mapped, the nucleic acid sequences may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP) (Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357).

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data (Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968). Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding KPP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation (Gatti, R.A. et al. (1988) Nature 336:577-580). The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, KPP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between KPP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest (Geysen, et al. (1984) PCT application WO84/03564). In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with KPP, or fragments thereof, and washed. Bound KPP is then detected by methods well known in the art. Purified KPP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding KPP specifically compete with a test compound for binding KPP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with KPP.

In additional embodiments, the nucleotide sequences which encode KPP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/340,235, U.S. Ser. No. 60/343,546, U.S. Ser. No. 60/343,007, U.S. Ser. No. 60/354,388, and U.S. Ser. No. 60/357,675, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Invitrogen), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Invitrogen), using the recommended procedures or similar methods known in the art (Ausubel et al., *supra*, ch. 5). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Biosciences) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Invitrogen, Carlsbad CA), PCDNA2.1 plasmid (Invitrogen), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Invitrogen.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Biosciences or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Amersham Biosciences); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard

ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (Ausubel et al., *supra*, ch. 7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) Nucleic Acids Res. 29:41-43); and HMM-based protein domain databases such as SMART (Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002) Nucleic Acids Res. 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families; see, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (MiraiBio, Alameda CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:43-84. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative kinases and phosphatases were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94; Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode kinases and phosphatases, the encoded polypeptides were analyzed by querying against PFAM models for kinases and phosphatases. Potential kinases and phosphatases were also identified by homology to Incyte cDNA sequences that had been annotated as kinases and phosphatases. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of KPP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:43-84 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:43-84 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook and Russell, *supra*, ch. 7; Ausubel et al., *supra*, ch. 4).

Analogous computer techniques applying BLAST were used to search for identical or related molecules in databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}\,\{\text{length(Seq. 1)}, \text{length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotides encoding KPP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding KPP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of KPP Encoding Polynucleotides

Full length polynucleotides are produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Biosciences), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Biosciences). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Biosciences), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Biosciences) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Biosciences) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotides are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Identification of Single Nucleotide Polymorphisms in KPP Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:43-84 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants.

An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezualan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:43-84 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Biosciences), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Biosciences). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing; see, e.g., Baldeschweiler et al., *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena, M., ed. (1999) DNA Microarrays: A Practical Approach, Oxford University Press, London). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements (Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31).

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Biosciences). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte Genomics). Specific control poly(A)$^+$ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Biosciences).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital

(A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte Genomics). Array elements that exhibit at least about a two-fold change in expression, a signal-to-background ratio of at least about 2.5, and an element spot size of at least about 40%, are considered to be differentially expressed.

XII. Complementary Polynucleotides

Sequences complementary to the KPP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring KPP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of KPP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the KPP-encoding transcript.

XIII. Expression of KPP

Expression and purification of KPP is achieved using bacterial or virus-based expression systems. For expression of KPP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express KPP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of KPP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding KPP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus (Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945).

In most expression systems, KPP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Biosciences). Following purification, the GST moiety can be proteolytically cleaved from KPP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). Purified KPP obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, XIX, XX, and XXI, where applicable.

XIV. Functional Assays

KPP function is assessed by expressing the sequences encoding KPP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT plasmid (Invitrogen, Carlsbad CA) and PCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994; Flow Cytometry, Oxford, New York NY).

The influence of KPP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding KPP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding KPP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of KPP Specific Antibodies

KPP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the KPP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art (Ausubel et al., *supra*, ch. 11).

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity (Ausubel et al., *supra*). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-KPP activity by, for example, binding the peptide or KPP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XVI. Purification of Naturally Occurring KPP Using Specific Antibodies

Naturally occurring or recombinant KPP is substantially purified by immunoaffinity chromatography using antibodies specific for KPP. An immunoaffinity column is constructed by covalently coupling anti-KPP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Biosciences). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing KPP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of KPP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/KPP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and KPP is collected.

XVII. Identification of Molecules Which Interact with KPP

KPP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent (Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled KPP, washed, and any wells with labeled KPP complex are assayed. Data obtained using different concentrations of KPP are used to calculate values for the number, affinity, and association of KPP with the candidate molecules.

Alternatively, molecules interacting with KPP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989; Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

KPP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVIII. Demonstration of KPP Activity

Generally, protein kinase activity is measured by quantifying the phosphorylation of a protein substrate by KPP in the presence of [γ-^{32}P]ATP. KPP is incubated with the protein substrate, ^{32}P-ATP, and an appropriate kinase buffer. The ^{32}P incorporated into the substrate is separated from free ^{32}P-ATP by electrophoresis and the incorporated ^{32}P is counted using a radioisotope counter. The amount of incorporated ^{32}P is proportional to the activity of KPP. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

In one alternative, protein kinase activity is measured by quantifying the transfer of gamma phosphate from adenosine triphosphate (ATP) to a serine, threonine or tyrosine residue in a protein substrate. The reaction occurs between a protein kinase sample with a biotinylated peptide substrate and gamma ^{32}P-ATP. Following the reaction, free avidin in solution is added for binding to the biotinylated ^{32}P-peptide product. The binding sample then undergoes a centrifugal ultrafiltration process with a membrane which will retain the product-avidin complex and allow passage of free gamma ^{32}P-ATP. The reservoir of the centrifuged unit containing the ^{32}P-peptide product as retentate is then counted in a scintillation counter. This procedure allows the assay of any type of protein kinase sample, depending on the peptide substrate and kinase reaction buffer selected. This assay is provided in kit form (ASUA, Affinity Ultrafiltration Separation Assay, Transbio Corporation, Baltimore MD, U.S. Patent No. 5,869,275). Suggested substrates and their respective enzymes include but are not limited to: Histone H1 (Sigma) and p34^{cdc2}kinase, Annexin I, Angiotensin (Sigma) and EGF receptor kinase, Annexin II and *src* kinase, ERK1 & ERK2 substrates and MEK, and myelin basic protein and ERK (Pearson, J.D. et al. (1991) Methods Enzymol. 200:62-81).

In another alternative, protein kinase activity of KPP is demonstrated in an assay containing KPP, 50 μl of kinase buffer, 1 μg substrate, such as myelin basic protein (MBP) or synthetic peptide substrates, 1 mM DTT, 10 μg ATP, and 0.5 μCi [γ-^{32}P]ATP. The reaction is incubated at 30°C for 30 minutes and stopped by pipetting onto P81 paper. The unincorporated [γ-^{32}P]ATP is removed by washing and the incorporated radioactivity is measured using a scintillation counter. Alternatively, the reaction is stopped by heating to 100°C in the presence of SDS loading buffer and resolved on a 12% SDS polyacrylamide gel followed by autoradiography. The amount of incorporated ^{32}P is proportional to the activity of KPP.

In yet another alternative, adenylate kinase or guanylate kinase activity of KPP may be measured by the incorporation of ^{32}P from [γ-^{32}P]ATP into ADP or GDP using a gamma radioisotope counter. KPP, in a kinase buffer, is incubated together with the appropriate nucleotide mono-phosphate substrate (AMP or GMP) and ^{32}P-labeled ATP as the phosphate donor. The reaction is incubated at 37°C and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected to gel electrophoresis to separate the mono-, di-, and triphosphonucleotide fractions. The diphosphonucleotide fraction is excised and counted. The radioactivity recovered is proportional to the activity of KPP.

In yet another alternative, other assays for KPP include scintillation proximity assays (SPA), scintillation plate technology and filter binding assays. Useful substrates include recombinant proteins tagged with glutathione transferase, or synthetic peptide substrates tagged with biotin. Inhibitors of KPP activity, such as small organic molecules, proteins or peptides, may be identified by such assays.

In another alternative, phosphatase activity of KPP is measured by the hydrolysis of para-nitrophenyl phosphate (PNPP). KPP is incubated together with PNPP in HEPES buffer pH 7.5, in the presence of 0.1% β-mercaptoethanol at 37°C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH (Diamond, R.H. et al. (1994) Mol. Cell. Biol. 14:3752-62). Alternatively, acid phosphatase activity of KPP is demonstrated by incubating KPP-containing extract with 100 μl of 10 mM PNPP in 0.1 M sodium citrate, pH 4.5, and 50 μl of 40 mM NaCl at 37°C for 20 min. The reaction is stopped by the addition of 0.5 ml of 0.4 M glycine/NaOH, pH 10.4 (Saftig, P. et al. (1997) J. Biol. Chem. 272:18628-18635). The increase in light absorbance at 410 nm resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the activity of KPP in the assay.

In the alternative, KPP activity is determined by measuring the amount of phosphate removed from a phosphorylated protein substrate. Reactions are performed with 2 or 4 nM KPP in a final volume of 30 μl containing 60 mM Tris, pH 7.6, 1 mM EDTA, 1 mM EGTA, 0.1% β-mercaptoethanol and 10 μM substrate, ^{32}P-labeled on serine/threonine or tyrosine, as appropriate. Reactions are initiated with substrate and incubated at 30° C for 10-15 min. Reactions are quenched with 450 μl of 4% (w/v) activated charcoal in 0.6 M HCl, 90 mM $Na_4P_2O_7$, and 2 mM NaH_2PO_4, then centrifuged at 12,000 × *g* for 5 min. Acid-soluble ^{32}Pi is quantified by liquid scintillation counting (Sinclair, C. et al. (1999) J. Biol. Chem. 274:23666-23672).

XIX. Kinase Binding Assay

Binding of KPP to a FLAG-CD44 cyt fusion protein can be determined by incubating KPP with anti-KPP-conjugated immunoaffinity beads followed by incubating portions of the beads (having 10-20 ng of protein) with 0.5 ml of a binding buffer (20 mM Tris-HCL (pH 7.4), 150 mM NaCl, 0.1% bovine serum albumin, and 0.05% Triton X-100) in the presence of ^{125}I-labeled FLAG-CD44cyt fusion protein (5,000 cpm/ng protein) at 4 °C for 5 hours. Following binding, beads were washed thoroughly in the binding buffer and the bead-bound radioactivity measured in a scintillation counter (Bourguignon, L.Y.W. et al. (2001) J. Biol. Chem. 276:7327-7336). The amount of incorporated ^{32}P is proportional to the amount of bound KPP.

XX. Identification of KPP Inhibitors

Compounds to be tested are arrayed in the wells of a 384-well plate in varying concentrations along with an appropriate buffer and substrate, as described in the assays in Example XVII. KPP activity is measured for each well and the ability of each compound to inhibit KPP activity can be determined, as well as the dose-response kinetics. This assay could also be used to identify molecules which enhance KPP activity.

XXI. Identification of KPP Substrates

A KPP "substrate-trapping" assay takes advantage of the increased substrate affinity that may be conferred by certain mutations in the PTP signature sequence of protein tyrosine phosphatases. KPP bearing these mutations form a stable complex with their substrate; this complex may be isolated biochemically. Site-directed mutagenesis of invariant residues in the PTP signature sequence in a clone encoding the catalytic domain of KPP is performed using a method standard in the art or a commercial kit, such as the MUTA-GENE kit from BIO-RAD. For expression of KPP mutants in *Escherichia coli*, DNA fragments containing the mutation are exchanged with the corresponding wild-type sequence in an expression vector bearing the sequence encoding KPP or a glutathione S-transferase (GST)-KPP fusion protein. KPP mutants are expressed in *E. coli* and purified by chromatography.

The expression vector is transfected into COS1 or 293 cells via calcium phosphate-mediated transfection with 20 μg of CsCl-purified DNA per 10-cm dish of cells or 8 μg per 6-cm dish. Forty-eight hours after transfection, cells are stimulated with 100 ng/ml epidermal growth factor to increase tyrosine phosphorylation in cells, as the tyrosine kinase EGFR is abundant in COS cells. Cells are lysed in 50 mM Tris·HCl, pH 7.5/5 mM EDTA/150 mM NaCl/1% Triton X-100/5 mM iodoacetic acid/10 mM sodium phosphate/10 mM NaF/5 μg/ml leupeptin/5 μg/ml aprotinin/1 mM benzamidine (1 ml per 10-cm dish, 0.5 ml per 6-cm dish). KPP is immunoprecipitated from lysates with an appropriate antibody. GST-KPP fusion proteins are precipitated with glutathione-Sepharose, 4 μg of mAb or 10 μl of beads respectively per mg of cell lysate. Complexes can be visualized by PAGE or further purified to identify substrate molecules (Flint, A.J. et al. (1997) Proc. Natl. Acad. Sci. USA 94:1680-1685).

Various modifications and variations of the described compositions, methods, and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. It will be appreciated that the invention provides novel and useful proteins, and their encoding polynucleotides, which can be used in the drug discovery process, as well as methods for using these compositions for the detection, diagnosis, and treatment of diseases and conditions. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Nor should the description of such embodiments be considered exhaustive or limit the invention to the precise forms disclosed. Furthermore, elements from one embodiment can be readily recombined with elements from one or more other embodiments. Such combinations can form a number of embodiments within the scope of the invention. It is intended that the scope of the invention be defined by the following claims and their equivalents.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7503679	1	7503679CD1	43	7503679CB1	7503679CB1
7503681	2	7503681CD1	44	7503681CB1	7503681CB1
7505819	3	7505819CD1	45	7505819CB1	7505819CB1
7505083	4	7505083CD1	46	7505083CB1	7505083CB1
7505866	5	7505866CD1	47	7505866CB1	7505866CB1
7503214	6	7503214CD1	48	7503214CB1	7503214CB1
7495312	7	7495312CD1	49	7495312CB1	7495312CB1
7506732	8	7506732CD1	50	7506732CB1	7506732CB1
7506736	9	7506736CD1	51	7506736CB1	7506736CB1
7507121	10	7507121CD1	52	7507121CB1	7507121CB1
90086258	11	90086258CD1	53	90086258CB1	90086258CB1
1967990	12	1967990CD1	54	1967990CB1	1967990CB1
3810039	13	3810039CD1	55	3810039CB1	3810039CB1
8032337	14	8032337CD1	56	8032337CB1	8032337CB1
7506411	15	7506411CD1	57	7506411CB1	7506411CB1
2658834	16	2658834CD1	58	2658834CB1	2658834CB1
6818489	17	6818489CD1	59	6818489CB1	6818489CB1
7509415	18	7509415CD1	60	7509415CB1	7509415CB1
7506916	19	7506916CD1	61	7506916CB1	7506916CB1
7507104	20	7507104CD1	62	7507104CB1	7507104CB1
7507105	21	7507105CD1	63	7507105CB1	7507105CB1
7507107	22	7507107CD1	64	7507107CB1	7507107CB1
7507109	23	7507109CD1	65	7507109CB1	7507109CB1
1833937	24	1833937CD1	66	1833937CB1	1833937CB1
7502036	25	7502036CD1	67	7502036CB1	7502036CB1
7503248	26	7503248CD1	68	7503248CB1	7503248CB1
7503968	27	7503968CD1	69	7503968CB1	7503968CB1
7505931	28	7505931CD1	70	7505931CB1	7505931CB1
7506912	29	7506912CD1	71	7506912CB1	7506912CB1
7506913	30	7506913CD1	72	7506913CB1	7506913CB1
7507029	31	7507029CD1	73	7507029CB1	7507029CB1
7507063	32	7507063CD1	74	7507063CB1	7507063CB1
7504755	33	7504755CD1	75	7504755CB1	7504755CB1
7509265	34	7509265CD1	76	7509265CB1	7509265CB1
7509371	35	7509371CD1	77	7509371CB1	7509371CB1
7509389	36	7509389CD1	78	7509389CB1	7509389CB1
7507005	37	7507005CD1	79	7507005CB1	7507005CB1
7509142	38	7509142CD1	80	7509142CB1	7509142CB1
7509157	39	7509157CD1	81	7509157CB1	7509157CB1
7509246	40	7509246CD1	82	7509246CB1	7509246CB1
7509380	41	7509380CD1	83	7509380CB1	7509380CB1
7509382	42	7509382CD1	84	7509382CB1	7509382CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	7503679CD1	g452194	1.9E-138	[Homo sapiens] protein tyrosine phosphatase (PTP-BAS, type 3) Maekawa,K., et al. (1994) FEBS Lett. 337:200-206 Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GLGF repeats
		321012|Ptpn13	8.6E-136	[Mus musculus][Protein phosphatase; Hydrolase][Cytoskeletal] Non-receptor protein tyrosine phosphatase with band 4.1 motif, may be involved in cell differentiation
		343730|PTPN13	1.3E-104	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Non-receptor protein tyrosine phosphatase that has similarity to cytoskeletal proteins and interacts with the Fas (TNFRSF6) cell-surface receptor, which mediates apoptosis Sato, T., et al. Science 268:411-5 (1995).
2	7503681CD1	g452194	1.8E-154	[Homo sapiens] protein tyrosine phosphatase (PTP-BAS, type 3) Maekawa,K., et al. (1994) FEBS Lett. 337:200-206 Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GLGF repeats
		343730|PTPN13	4.8E-155	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Non-receptor protein tyrosine phosphatase that has similarity to cytoskeletal proteins and interacts with the Fas (TNFRSF6) cell-surface receptor, which mediates apoptosis Sato, T., et al. Science 268:411-5 (1995).
		321012|Ptpn13	8.6E-153	[Mus musculus][Protein phosphatase; Hydrolase][Cytoskeletal] Non-receptor protein tyrosine phosphatase with band 4.1 motif, may be involved in cell differentiation
3	7505819CD1	g5225312	0.0	[Homo sapiens] calcineurin binding protein cabin 1 Sun,L., et al. (1998) Immunity 8:703-11 Cabin 1, a negative regulator for calcineurin signaling in T lymphocytes
		428820|KIAA0330	0.0	[Homo sapiens][Inhibitor or repressor] Protein with strong similarity to rat Rn.4 calcineurin inhibitor, which interacts with and non-competitively inhibits calcineurin activity, contains a TPR (tetratrico peptide repeat) domain, which may mediate protein-protein interactions Youn, H. D., and Liu, J. O. Immunity 13: 85-94 (2000).

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		327620|Cain	0.0	[Rattus norvegicus][Inhibitor or repressor] Calcineurin inhibitor, an inhibitor of calcineurin phosphatase activity, inhibits calcineurin-mediated synaptic vesicle endocytosis and cardiac hypertrophy, also may function in synaptic transmission and myoblast differentiation
4	7505083CD1	g2052058	6.0E-136	[Homo sapiens] SIRP-beta1 Kharitonenkov,A., et al (1997) Nature 38:181-186 A family of proteins that inhibit signalling through tyrosine kinase receptors
		343192|SIRP-BETA-1	5.2E-137	[Homo sapiens] Signal regulatory protein beta 1, associates with phosphorylated DAP12 (TYROBP), leading to efficient cell-surface expression, recruitment of tyrosine kinase SYK, and tyrosine phosphorylation of ERK1 (MAPK3)/ERK2 (MAPK1) and cell activation
5	7505866CD1	g4903263	0.0	[Homo sapiens] phosphorylase kinase alpha subunit Hendrickx,J., et al. (1999) Am. J. Hum. Genet. 64:1541-1549 Complete genomic structure and mutational spectrum of PHKA2 in patients with x-linked liver glycogenosis type I and II
		336932|PHKA1	0.0	[Homo sapiens][Regulatory subunit; Protein kinase; Transferase] Phosphorylase kinase regulatory subunit alpha-1 (muscle), which phosphorylates and thereby activates muscle-specific glycogen phosphorylase (PYGM); mutations in the corresponding gene are associated with muscle glycogenosis, a glycogen storage disease Wehner, M., et al. Hum Mol Genet 3:1983-7 (1994).
		368660|Phka1	0.0	[Mus musculus][Protein kinase; Regulatory subunit; Transferase] Phosphorylase kinase regulatory subunit alpha-1 (muscle), which phosphorylates and thereby activates muscle-specific glycogen phosphorylase (Pygm); mutations in the corresponding human gene are associated with several glycogen storage disorders
6	7503214CD1	g1679668	0.0	[Homo sapiens] mitogen-activated kinase kinase kinase 5 Wang,X.S., et al. J. Biol. Chem. 271:31607-31611 (1996)
		343076|MAP3K5	0.0	[Homo sapiens][Protein kinase; Transferase; Inhibitor or repressor] MAP kinase kinase kinase, activates SAPK/JNK and p38 signaling pathways, also induces apoptosis when overexpressed, may play a role in stress and cytokine-induced apoptosis Ichijo, H., et al. Science 275:90-4 (1997).

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		323210|Map3k5	0.0	[Mus musculus][Protein kinase; Transferase] Protein kinase that may play a role in tissue development during embryogenesis and cytokine-induced apoptosis; has strong similarity to human MEKK5, a MAP kinase kinase kinase (MAPKKK) that induces apoptosis when overexpressed Xia, Y., et al. Proc Natl Acad Sci U S A 97:5243-8 (2000).
		702206|nsy-1	1.7E-187	[Caenorhabditis elegans][Protein kinase] Serine/threonine protein kinase of the MAP kinase kinase kinase subfamily, putative ortholog of human MAP kinase kinase kinase (MAPKKK) ASK1 Sagasti, A., et al. Cell 105:221-32. (2001).
7	7495312CD1	g4768829	0.0	[Homo sapiens] putative protein-tyrosine kinase
		475681|LOC51086	0.0	[Homo sapiens][Protein kinase; Transferase] Protein containing ten ankyrin (Ank) repeats, which may mediate protein-protein interactions, and a eukaryotic protein kinase domain
		714105|C24A1.3	2.8E-90	[Caenorhabditis elegans][Protein kinase] Putative tyrosine kinase, has similarity to human mixed lineage kinases and D. melanogaster ABL protein kinase
8	7506732CD1	g1513315	2.6E-76	[Homo sapiens] guanylate kinase Brady,W.A., et al. J. Biol. Chem. 271:16734-16740 (1996)
		335704|GUK1	2.2E-77	[Homo sapiens][Transferase; Small molecule-binding protein] Guanylate kinase 1, catalyzes the conversion of GMP and GDP during GTP synthesis and the cGMP cycle, may function in phototransduction, involved in activation of antiviral drugs, may be a chemotherapy target; gene is downregulated in diffuse astrocytomas Huang, H., et al. Cancer Res 60:6868-74. (2000). Fitzgibbon, J., et al. FEBS Lett 385:185-8 (1996).
		583145|Guk1	2.6E-66	[Mus musculus][Transferase] Guanylate kinase 1, catalyzes the conversion of GMP and GDP during GTP synthesis and the cGMP cycle, involved in the activation of antiviral drugs and may be a chemotherapy target; the human GUK1 gene is downregulated in
9	7506736CD1	g1196436	4.4E-37	[Homo sapiens] guanylate kinase Fitzgibbon,J., et al. supra

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		335704|GUK1	3.9E-38	[Homo sapiens][Transferase; Small molecule-binding protein] Guanylate kinase 1, catalyzes the conversion of GMP and GDP during GTP synthesis and the cGMP cycle, may function in phototransduction, involved in activation of antiviral drugs, may be a chemotherapy target; gene is downregulated in diffuse astrocytomas Huang, H., et al. supra Brady, W. A., et al. J. Biol. Chem. 271:16734-40 (1996).
		583145|Guk1	1.4E-35	[Mus musculus][Transferase] Guanylate kinase 1, catalyzes the conversion of GMP and GDP during GTP synthesis and the cGMP cycle, involved in the activation of antiviral drugs and may be a chemotherapy target; the human GUK1 gene is downregulated in
10	7507121CD1	g5225312	0.0	[Homo sapiens] calcineurin binding protein cabin 1 Sun,L., et al. Immunity 8:703-11 (1998)
		428820|KIAA0330	0.0	[Homo sapiens][Inhibitor or repressor] Protein with strong similarity to rat Rn.4 calcineurin inhibitor, which interacts with and non-competitively inhibits calcineurin activity, contains a TPR (tetratricò peptide repeat) domain, which may mediate protein-protein interactions Youn, H. D., et al. Immunity 13:85-94 (2000).
		327620|Cain	0.0	[Rattus norvegicus][Inhibitor or repressor] Calcineurin inhibitor, an inhibitor of calcineurin phosphatase activity, inhibits calcineurin-mediated synaptic vesicle endocytosis and cardiac hypertrophy, also may function in synaptic transmission and myoblast differentiation Lai, M. M., et al. J. Biol. Chem. 275:34017-20 (2000).
11	90086258CD1	g2077825	2.1E-181	[Homo sapiens] MNK1 (Fukunaga, R. and Hunter, T. (1997) EMBO J. 16:1921-1933.)
		336418|MKNK1	1.9E-182	[Homo sapiens][Protein kinase; Transferase] MAP kinase-interacting serine/threonine kinase 1, a kinase activated by the ERK and p38 MAP kinase signaling pathways, phosphorylates translation initiation factor 4E (EIF4E) and reduces cap-dependent translation, phosphorylates cytosolic phospholipase A (Fukunaga, R., and Hunter, T. (1997) supra; Cuesta, R. et al. (2000) Genes And Development 14:1460-1470.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		618826|Mknk1	5.3E-169	[Mus musculus][Protein kinase; Transferase] MAP kinase-interacting serine/threonine kinase 1, a kinase that binds growth factor-regulated and stress-activated Erk and p38 MAP kinases and is activated by them, phosphorylates translation initiation factor 4E (Eif4e) (Waskiewicz, A. J. et al. (1997) EMBO J. 16:1909-1920.)
12	1967990CD1	g2564911	2.7E-277	[Homo sapiens] clk2 kinase (Winfield, S. L. et al. (1997) Genome Res. 7:1020-1026.)
		568158|CLK2	2.3E-278	[Homo sapiens][Protein kinase; Transferase] CDC-like kinase 2, a member of the LAMMER family of dual specificity protein kinases, associates with serine- and arginine-rich (SR) proteins and thus may regulate mRNA splicing, an alternative form lacks the kinase domain (Duncan, P. I. et al. (1998) Exp. Cell. Res. 241:300-308; Hanes, J. et al. (1994) J. Mol.
		325318|Clk2	1.3E-268	[Mus musculus][Protein kinase; Transferase] CDC-like kinase 2, a member of the LAMMER family of dual specificity protein kinases, may regulate mRNA splicing by phosphorylating serine- and arginine-rich (SR) proteins, an alternative form lacks the kinase domain
13	3810039CD1	g12005724	0.0	[Homo sapiens] mixed lineage kinase MLK1
		339748|MAP3K10	1.5E-239	[Homo sapiens][Protein kinase; Transferase] Member of the mixed-lineage kinase family, has SH3 and leucine zipper domains (Liu, Y. F. et al. (2000) J. Biol. Chem. 275:19035-19040.)
		321780|Map3k12	7.20E-61	[Mus musculus][Protein kinase; Transferase][Golgi; Nuclear; Cytoplasmic] Mitogen activated protein kinase kinase kinase 12, dual leucine zipper-bearing kinase, a member of the mixed lineage protein kinase (MLK), activates the SAPK/JNK and MAPK signaling pathways, may be involved in the regulation of cell growth and apoptosis (Xu, Z. et al. (2001) Mol. Cell. Biol. 21:4713-4724.)
14	8032337CD1	g4062852	3.7E-51	[Homo sapiens] p25 alpha.
		428310|p25	3.2E-52	[Homo sapiens] Activator of CDK5 which may contribute to Alzheimer's disease or other neurodegenerative diseases. Ahlijanian, M. K. et al. (2000) Proc. Natl. Acad. Sci. U. S. A. 97:2910-2915.

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
15	7506411CD1	g558436	0.0	[Homo sapiens] homolog of Drosophila discs large protein, isoform 2. Lue, R.A. et al. (1994) Proc. Natl. Acad. Sci. U.S.A. 91:9818-9822.
		340362|DLG1	0.0	[Homo sapiens][Adhesin/agglutinin Transferase; Other kinase][Basolateral plasma membrane; Cytoplasmic;Plasma membrane; Cell junction] Discs large homolog 1, a member of the membrane-associated guanylate kinase (MAGUK) family that has structural and signaling roles, localizes to regions of cell-cell contact, inhibits cell growth possibly through interaction with tumor suppressor proteins. Azim, A. C. et al. (1995) Genomics 30:613-616.
		583997|Dlgh1	0.0	[Mus musculus][Adhesin/agglutinin; Transferase; Other kinase][Basolateral plasma membrane; Axon; Plasma membrane; Cell junction] Discs large homolog 1, a member of the membrane-associated guanylate kinase (MAGUK) family that localizes to regions of cell-cell contact, has structural and signaling roles, plays a role in craniofacial development. Lue, R. A. et al. (1994) Proc. Natl. Acad. Sci. U. S. A. 91:9818-9822.
16	2658834CD1	g7248659	0.0	[Xenopus laevis] receptor protein tyrosine phosphatase LAR. Johnson, K.G. and Holt, C.E. (2000) Mech. Dev. 92:291-294.
		337418|PTPRF	0.0	[Homo sapiens][Protein phosphatase; Hydrolase; Receptor (signalling)][Plasma membrane] Protein tyrosine phosphatase receptor-type F (leukocyte common antigen related), regulates insulin receptor signaling and cell migration, inhibits tumor formation in nude mice. Yang, T. et al. (1999) Mol. Carcinog. 25:139-149.
		609837|Ptprf	0.0	[Rattus norvegicus][Protein phosphatase; Hydrolase Receptor (signalling)][Endoplasmic reticulum; Cytoplasmic; Plasma membrane] Protein tyrosine phosphatase receptor-type F (leukocyte common antigen related), regulates insulin receptor signaling, cell migration, and apoptosis; human PTPRF inhibits tumor formation in nude mice. Saito, S. et al. (1996) Biochem. Biophys. Res. Commun. 227:406-412.
17	6818489CD1	g3876094	2.8E-78	[Caenorhabditis elegans] Similarity to Aspergillus acid phosphatase (TR:G755244). Science (1998) 282:2012-2018.

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		243349|F21A3.2	4.5E-85	[Caenorhabditis elegans][Other phosphatase; Hydrolase] Member of the acid phosphatase protein family.
		243179|F18E2.1	2.4E-79	[Caenorhabditis elegans][Other phosphatase; Hydrolase] Member of the acid phosphatase protein family.
18	7509415CD1	g15341198	3.2E-141	[Mus musculus] Tau-tubulin kinase. Tomizawa, K. et al. (2001) FEBS letters. 492:221-227.
		249277|R90.1	2.6E-99	[Caenorhabditis elegans][Protein kinase; Transferase] Serine/threonine protein kinase of the casein kinase I subfamily, has similarity to S. cerevisiae Hrr25p casein kinase I.
		252302|Y38H8A.3	4.7E-58	[Caenorhabditis elegans][Protein kinase; Transferase] Serine/threonine protein kinase of the casein kinase I subfamily, has similarity to S. cerevisiae Yck1p and Yck2p casein kinase I isoforms.
19	7506916CD1	g7161864	4.2E-107	[Mus musculus] serine/threonine protein kinase Ruiz-Perez, V.L., et al. Nat. Genet. 24:283-286 (2000)
		599720|HSA250839	5.8E-108	[Homo sapiens][Protein kinase; Transferase] Protein containing a eukaryotic protein kinase domain Ruiz-Perez, V.L., et al. supra
		248325|M03C11.1	6.8E-80	[Caenorhabditis elegans][Protein kinase; Transferase] Serine/threonine protein kinase with similarity to human and D. melanogaster cAMP-dependent kinases, G protein-coupled receptor kinases and ribosomal protein S6 kinases
23	7507109CD1	441333|T23F11.1	1.7E-12	[Caenorhabditis elegans][Protein phosphatase; Hydrolase] Member of the protein phosphatase PP2C protein family
24	1833937CD1	g1477482	6.2E-19	[Homo sapiens] deoxyguanosine kinase Johansson, M. and Karlsson, A. Proc. Natl. Acad. Sci. U.S.A. 93:7258-7262 (1996)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		335036|DGUOK	3.4E-11	[Homo sapiens][Transferase; Other kinase][Cytoplasmic; Mitochondrial] Deoxyguanosine kinase, a purine salvage enzyme that is localized in the mitochondrial matrix and translocates to the cytosol during apoptosis Zhu, C., et al. J. Biol. Chem. 273:14707-11 (1998)
25	7502036CD1	g206152	2.9E-227	[Rattus norvegicus] calmodulin-dependent protein kinase II gamma subunit (EC 2.7.1.37) Tobimatsu, T., et al. J. Biol. Chem. 263:16082-16086 (1988)
		g21039158	0.0	[Homo sapiens] calcium/calmodulin-dependent protein kinase II gamma Gloyn, A.L. et al. (2002) Human calcium/calmodulin-dependent protein kinase II gamma gene (CAMK2G): cloning, genomic structure and detection of variants in subjects with Type II diabetes. Diabetologia 45:580-583
		331400|Rn.10961	2.4E-228	[Rattus norvegicus][Protein kinase; Transferase] Calcium/calmodulin-dependent protein kinase II gamma, activated by calmodulin binding and regulates Ca(2+)-mediated signaling pathways, may play a role in the developing and mature brain Sakagami, H., and Kondo, H. Brain Res. Mol. Brain Res. 20:51-63 (1993)
		618446|CAMK2B	7.9E-208	[Homo sapiens][Protein kinase; Transferase] Calcium calmodulin-dependent protein kinase II beta subunit, putative roles in signal transduction and cell growth, increased expression may play a role in schizophrenia; variant forms of the corresponding gene are expressed in tumor cells Williams, C. L., et al. Biochem. Pharmacol. 51:707-15 (1996)
26	7503248CD1	g5815141	3.7E-91	[Mus musculus] nuclear body associated kinase 1b
		692194|HIPK2	1.4E-90	[Homo sapiens][Protein kinase; Transferase; Inhibitor or repressor][Nuclear] Homeodomain interacting protein kinase 2, serine/threonine protein kinase, binds to and represses the transcriptional activity of homeodomain proteins, interaction with TRADD indicates a possible involvement in TNFRSF1A -mediated signaling Hofmann, T. G., et al. Biochimie 82:1123-7 (2000) Li, X., et al. Biochem. Biophys. Res. Commun. 277:513-7 (2000)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		584977|Hipk2	1.5E-90	[Mus musculus][Protein kinase; Transferase; Inhibitor or repressor][Nuclear] Homeodomain interacting protein kinase 2, protein kinase that binds to and represses the transcriptional activity of homeodomain proteins Kim, Y. H., et al. J. Biol. Chem. 273:25875-9 (1998)
27	7503968CD1	g15637110	9.0E-18	[Lycopersicon esculentum] LSTK-1-like kinase Pnueli, L., et al. Plant Cell 13:2687-2702 (2001)
		253122|ZC416.4	2.5E-17	[Caenorhabditis elegans][Protein kinase; Transferase] Serine/threonine protein kinase of the MAP kinase subfamily
		337214|MAPK9	2.1E-16	[Homo sapiens][Protein kinase; Transferase] Jun N-terminal kinase 2, a member of the MAP kinase family, regulates c-Jun (JUN) in response to proinflammatory cytokines or cellular stress; triggers apoptosis Potapova, O. et al. Mol. Cell. Biol. 20:1713-22 (2000)
28	7505931CD1	g5834426	2.2E-72	[Homo sapiens] glycerol kinase Sargent, C.A., et al. J. Med. Genet. 37:434-441 (2000)
		335526|GK	1.8E-73	[Homo sapiens][Transferase; Other kinase] Glycerol kinase, enzyme that catalyzes the phosphorylation of glycerol to glycerol 3-phosphate; deletion or mutation of the corresponding gene causes glycerol kinase deficiency which is characterized by hyperglyceroluria and hyperglycerolaemia Walker, A. P., et al. Am. J. Hum. Genet. 58:1205-11 (1996)
		583147|Gyk	8.9E-72	[Mus musculus][Transferase; Other kinase][Cytoplasmic; Mitochondrial outer membrane; Mitochondrial] Glycerol kinase, enzyme that catalyzes the phosphorylation of glycerol to glycerol 3-phosphate; deletion or mutation of the corresponding human GK gene causes glycerol kinase deficiency which is characterized by hyperglyceroluria and hyperglycerolaemia
29	7506912CD1	g6690020	3.3E-194	[Mus musculus] pantothenate kinase 1 beta Rock, C.O., et al. Pantothenate kinase regulation of the intracellular concentration of coenzyme A J. Biol. Chem. 275:1377-1383 (2000)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		690996|FLJ12899	4.9E-159	[Homo sapiens] Protein with low similarity to S. cerevisiae Ydr531p, which is a putative pantothenate kinase involved in coenzyme A biosynthesis
30	7506913CD1	g6690020	2.1E-165	[Mus musculus] pantothenate kinase 1 beta Rock, C.O., et al. supra
		690996|FLJ12899	1.3E-131	[Homo sapiens] Protein with low similarity to S. cerevisiae Ydr531p, which is a putative pantothenate kinase involved in coenzyme A biosynthesis
31	7507029CD1	g13183065	6.5E-67	[Mus musculus] dual-specificity phosphatase TS-DSP1
		428526|DUSP10	2.4E-14	[Homo sapiens][Protein phosphatase; Hydrolase][Nuclear; Cytoplasmic] Dual specificity phosphatase 10, a MAP kinase phosphatase that selectively dephosphorylates and inactivates the mitogen-activated protein kinases, and modulates MAP kinase-mediated signal transduction
		626536|Dusp10	2.4E-14	[Mus musculus][Protein phosphatase; Hydrolase] Dual specificity phosphatase 10, a putative MAP kinase phosphatase Masuda, K., et al. Cytogenet. Cell. Genet. 90:71-4 (2000)
32	7507063CD1	g1418936	5.4E-36	[Homo sapiens] protein-tyrosine-phosphatase Groom, L.A., et al. EMBO J. 15:3621-3632 (1996)
		347310|DUSP7	4.5E-37	[Homo sapiens][Protein phosphatase; Hydrolase] Dual specificity protein phosphatase-7, member of a subfamily of phosphatases that selectively dephosphorylates and inactivates mitogen-activated protein kinase, may be deleted or mutated in specific cancers Dowd, S., et al. J. Cell Sci. 111:3389-99 (1998)
		757458|Dusp6	2.6E-34	[Rattus norvegicus][Protein phosphatase; Hydrolase][Cytoplasmic] Mitogen activated protein kinase phosphatase 3 (dual specificity phosphatase 6), a cytosolic serine/threonine and tyrosine phosphatase, dephosphorylates and inactivates mitogen-activated protein kinase, induced in neurons by nerve growth factor Mourey, R. J., et al. J. Biol. Chem. 271:3795-802 (1996)
33	7504755CD1	g440855	3.3E-249	[Homo sapiens] leukocyte tyrosine kinase Toyoshima, H., et al. Proc. Natl. Acad. Sci. U.S.A. 90:5404-5408 (1993)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		336304|LTK	3.1E-148	[Homo sapiens][Protein kinase; Transferase; Receptor (signalling)][Plasma membrane] Leukocyte tyrosine kinase, a receptor tyrosine kinase of the insulin receptor family; exists in alternatively spliced forms with variable extracellular domains that differentially localize within the cell Ueno, H., et al. J. Biol. Chem. 270:20135-42 (1995)
		339978|ALK	1.3E-100	[Homo sapiens][Protein kinase; Transferase; Receptor (signalling)][Plasma membrane] Anaplastic lymphoma kinase, a receptor protein tyrosine kinase that may be involved in brain development; fusion of the gene to the nucleophosmin (NPM) gene via chromosomal translocation is associated with non-Hodgkin's lymphomas Souttou, B., et al. J. Biol. Chem. 276:9526-31 (2001) Lawrence, B., et al. Am. J. Pathol. 157:377-84 (2000)
		583579|Alk	2.1E-99	[Mus musculus][Protein kinase; Transferase; Receptor (signalling)][Plasma membrane] Anaplastic lymphoma kinase, a receptor protein tyrosine kinase that may be involved in the development of the central nervous system; fusion of the human ALK gene to the human nucleophosmin (human NPM) gene is associated with non-Hodgkin's lymphomas Slupianek, A., et al. Cancer Res. 61:2194-9. (2001)
34	7509265CD1	g12803641	9.8E-89	[Homo sapiens] Similar to cell cycle related kinase
		g23344742	7.0E-88	[Homo sapiens] cell cycle related kinase
		568698|CCRK	2.6E-82	[Homo sapiens][Protein kinase; Transferase] Protein containing a kinase domain, has a region of moderate similarity to a region of cyclin-dependent kinases, which are involved in cell cycle control
35	7509371CD1	g15825377	3.6E-142	[Mus musculus] NIMA-related kinase 8
		253196|ZC581.1	2.2E-60	[Caenorhabditis elegans][Protein kinase; Transferase] Serine/threonine protein kinase with similarity to human NIMA-related kinases and D. melanogaster and S. cerevisiae calcium/calmodulin kinases
		338322|STK2	8.4E-50	[Homo sapiens][Protein kinase; Transferase] Serine/threonine kinase that is most highly expressed in the heart Cance, W. G., et al. Int. J. Cancer 54:571-7 (1993)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		430068|Nek4	2.0E-49	[Mus musculus][Protein kinase; Transferase] NIMA (never in mitosis gene a)-related expressed kinase 4, member of NIMA kinase family which controls entrance into mitosis, and may play a role as a cell cycle regulator, highly expressed in testis Chen, A., et al. Gene 234:127-37 (1999)
36	7509389CD1	g15825377	1.7E-57	[Mus musculus] NIMA-related kinase 8
		253196|ZC581.1	3.6E-28	[Caenorhabditis elegans][Protein kinase; Transferase] Serine/threonine protein kinase with similarity to human NIMA-related kinases and D. melanogaster and S. cerevisiae calcium/calmodulin kinases
		338322|STK2	1.3E-21	[Homo sapiens][Protein kinase; Transferase] Serine/threonine kinase that is most highly expressed in the heart Cance, W. G., et al. supra
		430068|Nek4	8.0E-21	[Mus musculus][Protein kinase; Transferase] NIMA (never in mitosis gene a)-related expressed kinase 4, member of NIMA kinase family which controls entrance into mitosis, and may play a role as a cell cycle regulator, highly expressed in testis Chen, A., et al. supra
37	7507005CD1	g3218467	3.6E-85	[Gallus gallus] putative phosphatase Houston, B., et al. Biochim. Biophys. Acta 1480:500-506 (1999)
		g20196839	1.0E-147	[Homo sapiens]phosphatase, orphan 1
38	7509142CD1	g2665458	3.0E-42	[Mus musculus] protein-tyrosine-phosphatase Ohsugi, M., et al. J. Biol. Chem. 272:33092-9 (1997)
		425672|DKFZP566K0524	1.5E-79	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Protein with high similarity to murine Mm.8019, which is a cytoplasmic protein tyrosine phosphatase that is specifically expressed in testicular germ cells, member of the protein-tyrosine phosphatase family
		582661|Ptpn20	2.5E-43	[Mus musculus][Protein phosphatase; Hydrolase][Cytoplasmic] Protein tyrosine phosphatase non-receptor type 20, a testis-specific protein tyrosine phosphatase that may play a role in spermatogenesis or meiosis Ohsugi, M., et al. supra
39	7509157CD1	g2665458	6.0E-108	[Mus musculus] protein-tyrosine-phosphatase Ohsugi, M., et al. supra

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		425672|DKFZP566K0524	2.2E-140	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Protein with high similarity to murine Mm.8019, which is a cytoplasmic protein tyrosine phosphatase that is specifically expressed in testicular germ cells, member of the protein-tyrosine phosphatase family
		582661|Ptpn20	5.1E-109	[Mus musculus][Protein phosphatase; Hydrolase][Cytoplasmic] Protein tyrosine phosphatase non-receptor type 20, a testis-specific protein tyrosine phosphatase that may play a role in spermatogenesis or meiosis Ohsugi, M., et al. supra
40	7509246CD1	g1418936	5.4E-36	[Homo sapiens] protein-tyrosine-phosphatase Groom, L.A., et al. supra
		347310|DUSP7	4.5E-37	[Homo sapiens][Protein phosphatase; Hydrolase] Dual specificity protein phosphatase-7, member of a subfamily of phosphatases that selectively dephosphorylates and inactivates mitogen-activated protein kinase, may be deleted or mutated in specific cancers Dowd, S., et al. supra
		662410|DUSP6	2.9E-35	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Dual specificity phosphatase 6, a cytosolic phosphatase that selectively dephosphorylates and inactivates mitogen-activated protein kinases, downregulated in some pancreatic cancer cell lines Rossig, L., et al. J. Biol. Chem. 275:25502-7 (2000)
		757458|Dusp6	2.6E-34	[Rattus norvegicus][Protein phosphatase; Hydrolase][Cytoplasmic] Mitogen activated protein kinase phosphatase 3 (dual specificity phosphatase 6), a cytosolic serine/threonine and tyrosine phosphatase, dephosphorylates and inactivates mitogen-activated protein kinase, induced in neurons by nerve growth factor Mourey, R. J., et al. supra
41	7509380CD1	g2665458	4.5E-135	[Mus musculus] protein-tyrosine-phosphatase Ohsugi, M., et al. supra
		425672|DKFZP566K0524	4.1E-194	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Protein with high similarity to murine Mm.8019, which is a cytoplasmic protein tyrosine phosphatase that is specifically expressed in testicular germ cells, member of the protein-tyrosine phosphatase family

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		582661|Ptpn20	3.8E-136	[Mus musculus][Protein phosphatase; Hydrolase][Cytoplasmic] Protein tyrosine phosphatase non-receptor type 20, a testis-specific protein tyrosine phosphatase that may play a role in spermatogenesis or meiosis Ohsugi, M., et al. supra
42	7509382CD1	g2665458	1.1E-42	[Mus musculus] protein-tyrosine-phosphatase Ohsugi, M., et al. supra
		425672|DKFZP566K0524	5.5E-80	[Homo sapiens][Protein phosphatase; Hydrolase][Cytoplasmic] Protein with high similarity to murine Mm.8019, which is a cytoplasmic protein tyrosine phosphatase that is specifically expressed in testicular germ cells, member of the protein-tyrosine phosphatase family
		582661|Ptpn20	9.6E-44	[Mus musculus][Protein phosphatase; Hydrolase][Cytoplasmic] Protein tyrosine phosphatase non-receptor type 20, a testis-specific protein tyrosine phosphatase that may play a role in spermatogenesis or meiosis Ohsugi, M., et al. supra

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	7503679CD1	1179	signal_cleavage: M1-G23	SPSCAN
			Signal Peptide: M1-A18, M1-G23, M8-A25	HMMER
			FERM domain (Band 4.1 family): T399-H526	HMMER_PFAM
			PDZ domain (Also known as DHR or GLGF): F820-R904, E949-P1036	HMMER_PFAM
			Band 4.1 family domain proteins BL00660: G327-C379, R420-P459, D506-I549, F564-G587, F592-N614	BLIMPS_BLOCKS
			Band 4.1 family domain signatures and profile: A501-E552	PROFILESCAN
			Band 4.1 protein family signature PR00935: A351-Y363, L425-C438, C438-Y458, D506-G522	BLIMPS_PRINTS
			PDZ domain proteins PF00595: I996-N1006	BLIMPS_PFAM
			PROTEIN SH3 DOMAIN REPEAT PD00289: G999-G1012	BLIMPS_PRODOM
			PROTEIN CYTOSKELETON STRUCTURAL PHOSPHATASE HYDROLASE PROTEIN TYROSINE PHOSPHORYLATION MOESIN TYROSINE BAND PD000961: L320-V523	BLAST_PRODOM
			PHOSPHATASE TYROSINE PROTEIN TYPE PTP BAS HYDROLASE PROTEIN TYROSINE PHOSPHATASE PHOSPHOTYROSINE PTP-ASE 1E PD150192: H526-Q621	BLAST_PRODOM
			BAND 4.1 DM00609 A54971|562-990: T308-Q621, I652-F663; S51005|13-453: T308-F613; JC4155|11-447: K312-Q616	BLAST_DOMO
			GLGF DOMAIN DM00224|A54971|1358-1454: S809-E902	BLAST_DOMO
			Potential Phosphorylation Sites: S27 S75 S92 S129 S164 S211 S219 S256 S262 S273 S285 S478 S530 S596 S656 S658 S751 S809 S814 S914 S918 S941 S1078 S1098 S1155 T115 T126 T162 T192 T308 T569 T580 T586 T591 T599 T677 T701 T882 T1063 T1143 Y437 Y468	MOTIFS
			Potential Glycosylation Sites: N48 N180 N555 N743 N839 N916	MOTIFS
2	7503681CD1	1170	FERM domain (Band 4.1 family): T284-H411	HMMER_PFAM
			PDZ domain (Also known as DHR or GLGF): F811-R895, R636-S721, E940-P1027	HMMER_PFAM
			Band 4.1 family domain proteins BL00660: R305-P344, D391-I434, F449-G472, F477-N499, G212-C264	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Band 4.1 family domain signatures and profile: A386-E437	PROFILESCAN
			Band 4.1 protein family signature PR00935: A236-Y248, L310-C323, C323-Y343, D391-G407	BLIMPS_PRINTS
			PDZ domain proteins PF00595: I987-N997	BLIMPS_PFAM
			PROTEIN SH3 DOMAIN REPEAT PD00289: G990-G1003	BLIMPS_PRODOM
			PROTEIN CYTOSKELETON STRUCTURAL PHOSPHATASE HYDROLASE PROTEIN TYROSINE PHOSPHORYLATION MOESIN TYROSINE BAND PD000961: L205-V408	BLAST_PRODOM
			PHOSPHATASE TYROSINE PROTEIN TYPE PTP BAS HYDROLASE PROTEIN TYROSINE PHOSPHATASE PHOSPHOTYROSINE PTP-ASE 1E PD150192: H411-Q506	BLAST_PRODOM
			BAND 4.1 DM00609 A54971|562-990: T193-Q506, I537-F548; S51005|13-453: T193-F498; JC4155|11-447: K197-Q501	BLAST_DOMO
			GLGF DOMAIN DM00224|A54971|1358-1454: S800-E893	BLAST_DOMO
			Potential Phosphorylation Sites: S14 S49 S96 S104 S141 S147 S158 S170 S363 S415 S481 S541 S543 S625 S742 S800 S805 S905 S909 S932 S1069 S1089 S1146 T11 T47 T77 T193 T454 T465 T471 T476 T484 T562 T586 T638 T680 T873 T1054 T1134 Y322 Y353	MOTIFS
			Potential Glycosylation Sites: N65 N440 N734 N830 N907	MOTIFS
3	7505819CD1	1977	CAIN PD186007: M1-I1977	BLAST_PRODOM
			Cell attachment sequence: R1269-D1271	MOTIFS
			N-4 cytosine-specific DNA methylases signature: L1131-Y1136	MOTIFS
			Potential Phosphorylation Sites: S8 S15 S154 S159 S169 S180 S252 S303 S308 S320 S333 S353 S387 S462 S478 S505 S518 S713 S741 S781 S968 S1097 S1174 S1184 S1189 S1192 S1301 S1457 S1477 S1481 S1538 S1557 S1584 S1591 S1630 S1641 S1772 S1831 S1857 S1916 S1920 S1946 S1958 S1962 S1966 T55 T71 T175 T248 T680 T714 T727 T775 T805 T822 T972 T1093 T1139 T1156 T1194 T1208 T1262 T1328 T1442 T1449 T1713 T1732 T1846 T1968 Y89 Y491 Y782	MOTIFS
			Potential Glycosylation Sites: N264 N457 N925 N1362 N1389	MOTIFS
4	7505083CD1	279	signal_cleavage: M1-A26	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Immunoglobulin domain: E43-T92, E146-Y190, R204-S222	HMMER_PFAM
			Cytosolic domain: K276-A279; Transmembrane domain: L253-W275; Non-cytosolic domain: M1-L252	TMHMMER
			Immunoglobulins and major histocompatibility complex proteins signature: D187-H242	PROFILESCAN
			SIGNAL REGULATORY PROTEIN BETA1 PRECURSOR SIRP BETA1 SIGNAL IMMUNOGLOBULIN FOLD GLYCOPROTEIN TRANSMEMBRANE PD054160: Q218-A279	BLAST_PRODOM
			Cell attachment sequence: R99-D101	MOTIFS
			Potential Phosphorylation Sites: S56 S174 T25 T61 T76 T92 T98 T166 T181	MOTIFS
			Potential Glycosylation Sites: N125 N150 N172	MOTIFS
5	7505866CD1	1207	PHOSPHORYLASE KINASE ALP PD01841: K73-G122, F144-W184, E185-A205, K206-A257, G258-R295, F296-D344, Q349-W402, K435-I471, L494-D535, Q892-L928, W1039-H1068, T1081-L1122, F1172-A1190	BLIMPS_PRODOM
			PHOSPHORYLASE B KINASE REGULATORY CHAIN SUBUNIT GLYCOGEN METABOLISM PHOSPHORYLATION CALMODULIN BINDING PD005098: R10-R428, H726-G965, K397-S642, I478-L494	BLAST_PRODOM
			PHOSPHORYLASE B KINASE REGULATORY CHAIN SUBUNIT GLYCOGEN METABOLISM PHOSPHORYLATION CALMODULIN BINDING PD008424: D1024-Q1151, H1068-Q1207	BLAST_PRODOM
			PHOSPHORYLASE B KINASE ALPHA REGULATORY CHAIN LIVER ISOFORM L SUBUNIT	BLAST_PRODOM
			PHOSPHORYLASE B KINASE ALPHA REGULATORY CHAIN ISOFORM SUBUNIT GLYCOGEN METABOLISM PD014573: V966-Q1009	BLAST_PRODOM
			CALMODULIN-BINDING DOMAIN DM03490 S24109|1-1235: M1-Q1009, R1010-Q1205; P46020|1-1222: M1-Q1207; P34335|61-1256: M1-E633, V836-V1191, S580-L837; P12798|36-1091: R10-R428, G396-S642, S735-G965, E1034-E1197	BLAST_DOMO
			Leucine zipper pattern: L830-L851	MOTIFS
			Cell attachment sequence: R186-D188	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Phosphopantetheine attachment site: E196-A211	MOTIFS
			Neutral zinc metallopeptidases, zinc-binding region signature: I982-G991	MOTIFS
			Potential Phosphorylation Sites: S107 S431 S433 S464 S495 S640 S656 S672 S695 S881 S935 S947 S960 S969 S983 S1021 S1123 T158 T251 T279 T303 T507 T557 T585 T617 T993 T995 T1081 T1086 Y593	MOTIFS
			Potential Glycosylation Sites: N198 N277 N493 N603 N630 N794 N933	MOTIFS
6	7503214CD1	1328	Protein kinase domain: V672-L923	HMMER_PFAM
			Protein kinases signatures and profile: T764-G816	PROFILESCAN
			Tyrosine kinase catalytic domain signature PR00109: M741-R754, Y778-V796, G826-I836, A847-L869, L892-T914	BLIMPS_PRINTS
			KINASE APOPTOSIS ASK1 PROTEIN MEK SIGNAL REGULATING MITOGEN-ACTIVATED MEKK5 MAP/ERK MAPKKK5 PD018410: V75-N620	BLAST_PRODOM
			KINASE PROTEIN APOPTOSIS ASK1 MEK SIGNAL REGULATING MITOGEN-ACTIVATED MEKK5 MAP/ERK MAPKKK5 PD014104: P950-G1173	BLAST_PRODOM
			KINASE APOPTOSIS ASK1 MEK SIGNAL REGULATING MITOGEN-ACTIVATED MEKK5 MAP/ERK MAPKKK5 MITOGEN PD024456: E1183-R1316	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|A48084|98-348: V672-R911; DM00004|Q01389|1176-1430: V672-T913; DM00004|Q10407|826-1084: V672-T913; DM00004|P41892|11-249: L673-T914	BLAST_DOMO
			Protein kinases ATP-binding region signature: L673-K696	MOTIFS
			Serine/Threonine protein kinases active-site signature: I784-V796	MOTIFS
			Potential Phosphorylation Sites: S23 S56 S212 S253 S338 S382 S432 S486 S550 S609 S625 S645 S730 S811 S902 S959 S993 S999 S1008 S1009 S1024 S1052 T48 T205 T218 T428 T466 T545 T653 T764 T810 T855 T861 T913 T951 T1202 T1255 T1282 T1291 Y778 Y1281	MOTIFS
			Potential Glycosylation Sites: N381 N620	MOTIFS
7	7495312CD1	603	Ankyrin repeat: D381-C413, G133-A165, V166-E198, D339-C371, N100-Y132, V199-N233, Y269-N302, E234-I268, F304-R338, N66-R99	HMMER_PFAM
			Protein kinase domain: I463-L589	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Tyrosine kinase catalytic domain signature PR00109: T539-H552	BLIMPS_PRINTS
			Ank repeat proteins. PF00023: G270-Y279	BLIMPS_PFAM
			REPEAT PROTEIN ANK NUCLE PD00078: N267-Y279	BLIMPS_PRODOM
			ATP-BINDING TRANSFERASE C24A1.3 PROTEIN PD083398: D381-I463	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|Q05609|553-797: E467-L589; DM00004|P18160|1291-1551: F465-N590; DM00004|S52578|118-367: F465-L589; DM00004|I38044|100-349: E467-L589	BLAST_DOMO
			Protein kinases ATP-binding region signature: I469-K490	MOTIFS
			Potential Phosphorylation Sites: S21 S123 S208 S224 S257 S293 S376 S461 S499 S501 T11	MOTIFS
			Potential Glycosylation Sites: N194 N579	MOTIFS
8	7506732CD1	160	Guanylate kinase: R23-S107	HMMER_PFAM
			Guanylate kinase proteins BL00856: E27-K74	BLIMPS_BLOCKS
			GUANYLATE KINASE GMP TRANSFERASE ATP-BINDING ACETYLATION PD003452: L108-G159	BLAST_PRODOM
			PROTEIN DOMAIN SH3 KINASE GUANYLATE TRANSFERASE ATP-BINDING REPEAT GMP MEMBRANE PD001338: N42-S107	BLAST_PRODOM
			GUANYLATE KINASE DM00755|P46195|1-194: S2-N42, N42-G159; DM00755|S32545|1-196: P4-N42, N42-G159; DM00755|P15454|1-185: R5-R41, N42-E151; DM00755|P21074|1-192: M1-R41, G33-K153	BLAST_DOMO
			ATP/GTP-binding site motif A (P-loop): G11-S18	MOTIFS
			Potential Phosphorylation Sites: S121 T39 T46 T102	MOTIFS
			Potential Glycosylation Sites: N134	MOTIFS
9	7506736CD1	102	Guanylate kinase: T40-S102	HMMER_PFAM
			Guanylate kinase proteins BL00856: V8-L20, V36-V56	BLIMPS_BLOCKS
			GUANYLATE KINASE DM00755|P46195|1-194: S2-S77; DM00755|S32545|1-196: P4-S77; DM00755|P15454|1-185: R5-S77; DM00755|P21074|1-192: M1-F76	BLAST_DOMO
			ATP/GTP-binding site motif A (P-loop): G11-S18	MOTIFS
			Guanylate kinase signature: T39-V56	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Phosphorylation Sites: S77 T39	MOTIFS
10	7507121CD1	2191	TPR Domain: Y90-D123, V124-H157, V615-S648, N1055-R1088	HMMER_PFAM
			CAIN PD186007: C189-I2191	BLAST_PRODOM
			Leucine zipper pattern: L164-L185	MOTIFS
			Potential Phosphorylation Sites: S11 S20 S23 S55 S73 S74 S212 S222 S229 S368 S373 S383 S394 S466 S517 S522 S534 S547 S567 S601 S676 S692 S719 S732 S927 S955 S995 S1182 S1311 S1388 S1398 S1403 S1406 S1515 S1671 S1691 S1695 S1752 S1771 S1798 S1805 S1844 S1855 S1986 S2045 S2071 S2130 S2134 S2160 S2172 S2176 S2180	MOTIFS
			Potential Phosphorylation Sites: T12 T92 T108 T170 T269 T285 T389 T462 T894 T928 T941 T989 T1019 T1036 T1186 T1307 T1353 T1370 T1408 T1422 T1476 T1542 T1656 T1663 T1927 T1946 T2060 T2182 Y303 Y705 Y996	MOTIFS
			Potential Glycosylation Sites: N8 N478 N671 N1139 N1576 N1603	MOTIFS
11	90086258CD1	347	Protein kinase domain: K49-V333	HMMER_PFAM
			Protein kinases signatures and profile E146-N201	PROFILESCAN
			Tyrosine kinase catalytic domain signature PR00109:F124-Q137, F160-C178, G217-V227, C241-H263, I302-A324	BLIMPS_PRINTS
			MNK1 PD047052: M1-M47	BLAST_PRODOM
			KINASE PROTEIN TRANSFERASE ATP-BINDING SERINE/THREONINE PROTEIN PHOSPHORYLATION RECEPTOR TYROSINE PROTEIN PRECURSOR TRANSMEMBRANE PD000001: T213-H339, Y99-S203	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|A57459|417-662: E53-F260, Y237-A324; DM00004|A53300|420-665: E53-F260, Y237-A324; DM00004|P18653|409-654: E53-F260, Y237-A324; DM00004|S57347|21-266: E53-F260, Y237-S323	BLAST_DOMO
			Protein kinases ATP-binding region signature: L55-K78	MOTIFS
			Serine/Threonine protein kinases active-site signature: I166-C178	MOTIFS
			Potential Phosphorylation Sites: S4 S286 T208 T235 Y99	MOTIFS
			Potential Glycosylation Sites: N201	MOTIFS
12	1967990CD1	498	Protein kinase domain: Y162-F478	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Protein kinases signatures and profile: Q265-H336	PROFILESCAN
			PROTEIN KINASE CLK2 TRANSFERASE SERINE/THREONINE-PROTEIN ATP-BINDING TYROSINE-PROTEIN PHOSPHORYLATION NUCLEAR ALTERNATIVE PD027189: E48-Y107	BLAST_PRODOM
			KINASE TRANSFERASE PROTEIN SERINE/THREONINE-PROTEIN ATPBINDING II PHOSPHORYLATION CASEIN ALPHA CHAIN PD002608: H334-A479	BLAST_PRODOM
			PROTEIN KINASE TRANSFERASE SERINE/THREONINE-PROTEIN ATPBINDING TYROSINE-PROTEIN PHOSPHORYLATION NUCLEAR CLK1 CLK PD016041: M1-R161	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|P49760|165-469: I164-L469; DM00004|P49761|158-462: I164-L469; DM00004|P49762|172-474: I164-L469; DM00004|P49759|163-467: I164-L469	BLAST_DOMO
			Cell attachment sequence: R85-D87	MOTIFS
			Crystallins beta and gamma 'Greek key' motif signature: V226-S241	MOTIFS
			Protein kinases ATP-binding region signature: L168-K192	MOTIFS
			Serine/Threonine protein kinases active-site signature: L285-F297	MOTIFS
			Potential Phosphorylation Sites: S10 S17 S24 S32 S37 S39 S57 S63 S75 S106 S109 S112 S113 S124 S135 S136 S141 S248 S300 S316 S342 S443 S457 S492 T42 T167 T286 T407 T468 Y51 Y68 Y72 Y90	MOTIFS
			Potential Glycosylation Sites: N104 N423	MOTIFS
13	3810039CD1	1081	signal_cleavage: M1-A17	SPSCAN
			Signal Peptide: M1-A16, M1-A18, M1-P20, M1-D23, M1-G24, M1-A27	HMMER
			SH3 domain: P55-R114	HMMER_PFAM
			Protein kinase domain: L144-L403	HMMER_PFAM
			Receptor tyrosine kinase class II proteins BL00239: E191-P238, L242-I264, W301-R350, L355-I399	BLIMPS_BLOCKS
			Receptor tyrosine kinase class III proteins BL00240: D180-G234, E300-V347, V347-I399	BLIMPS_BLOCKS
			Receptor tyrosine kinase class V proteins BL00790: R114-G153, I164-C217, S308-W340, L366-M414	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Protein kinases signatures and profile: L242-T304	PROFILESCAN
			Tyrosine kinase catalytic domain signature PR00109: M220-S233, D258-I276, G311-I321, S330-I352, C374-F396	BLIMPS_PRINTS
			SH3 domain signature PR00452: P55-A65, D69-K84, D91-N100, R102-R114	BLIMPS_PRINTS
			KINASE DOMAIN SH3 MIXED LINEAGE SERINE/THREONINE WITH LEUCINE ZIPPER PROLINE PD024997: I406-A749, F419-E833	BLAST_PRODOM
			KINASE PROTEIN TRANSFERASE ATP-BINDING SERINE/THREONINE PROTEIN PHOSPHORYLATION RECEPTOR TYROSINE PROTEIN PRECURSOR TRANSMEMBRANE PD000001: L242-A317, L146-F222, W315-F349	BLAST_PRODOM
			DOMAIN KINASE SH3 MIXED LINEAGE SERINE/THREONINE WITH LEUCINE ZIPPER PROLINE PD034700: N855-R950, S888-P1006	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|A53800|119-368: L146-F396; DM00004|I38044|100-349: L146-F396; DM00004|JC2363|126-356: W163-F396, I150-I276	BLAST_DOMO
			ZIPPER MOTIF LEUCINE DM08113|I38044|392-721: R438-A749, Q436-P681, P869-P893	BLAST_DOMO
			Protein kinases ATP-binding region signature: I150-K171	MOTIFS
			Serine/Threonine protein kinases active-site signature: I264-I276	MOTIFS
			Potential Phosphorylation Sites: S89 S118 S233 S286 S541 S569 S611 S618 S648 S715 S778 S789 S816 S822 S829 S842 S888 S958 S1001 S1007 S1018 T72 T112 T145 T304 T373 T404 T405 T446 T565 T785 T892 T948 T954 T1019 T1040 T1043 T1067 Y335	MOTIFS
			Potential Glycosylation Sites: N821 N870 N999	MOTIFS
14	8032337CD1	170	Rhodanese proteins BL00380: F30-V45	BLIMPS_BLOCKS
			Muscarinic M2 receptor signature PR00539: D41-S55	BLIMPS_PRINTS
			BRAIN SPECIFIC PROTEIN P25 CODED FOR BY C ELEGANS cDNA YK112E11.5 PD035774: F9-K167	BLAST_PRODOM
			Potential Phosphorylation Sites: S3 S21 S86 S137 T46 T48 T123 T129 Y126	MOTIFS
			Potential Glycosylation Sites: N29	MOTIFS
15	7506411CD1	893	Guanylate kinase: T736-Q838	HMMER_PFAM
			PDZ domain (Also known as DHR or GLGF).: E191-R277, E286-P372, K433-R513	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Guanylate kinase proteins BL00856: V732-V752, D761-Q808, Y189-L201	BLIMPS_BLOCKS
			SH3 domain signature PR00452: L551-K561, Q570-A585, D587-V596, E604-K616	BLIMPS_PRINTS
			PDZ domain proteins (Als PF00595: L473-N483	BLIMPS_PFAM
			PROTEIN SH3 DOMAIN REPEAT PD00289: G476-A489	BLIMPS_PRODOM
			HOMOLOG DISCS LARGE PRESYNAPTIC PROTEIN SAP97 SYNAPSE-ASSOCIATED SH3 DOMAIN REPEAT PD015466: M1-P98	BLAST_PRODOM
			HOMOLOG PROTEIN SH3 DOMAIN REPEAT DISCS LARGE PRESYNAPTIC SAP97 SYNAPSE-ASSOCIATED PD012559: S100-E190	BLAST_PRODOM
			PROTEIN DOMAIN SH3 KINASE GUANYLATE TRANSFERASE ATP-BINDING REPEAT GMP MEMBRANE PD001338: T735-Q838	BLAST_PRODOM
			PROTEIN SH3 DOMAIN REPEAT PRESYNAPTIC SYNAPSE ASSOCIATED HOMOLOG DENSITY ALTERNATIVE SPLICING PD005212: K643-T735	BLAST_PRODOM
			GUANYLATE KINASE DM00755|I38757|709-898: Q698-P888; |P31007|765-954: V700-P888; |P31016|529-718: E699-P888	BLAST_DOMO
			GLGF DOMAIN DM00224|I38757|213-307: V180-I508	BLAST_DOMO
			Guanylate kinase signature: T735-V752	MOTIFS
			Potential Phosphorylation Sites: S23 S32 S138 S143 S281 S518 S542 S546 S586 S610 S637 S648 S660 S665 S676 S681 S784 T8 T59 T183 T350 T624 T735 T753 T835 T865 Y366 Y773	MOTIFS
			Potential Glycosylation Sites: N141 N181 N360 N388 N533 N584 N674 N701	MOTIFS
16	2658834CD1	1796	signal_cleavage: M1-G29	SPSCAN
			Signal Peptide: M11-G26, M11-G29, M1-G29, M11-S31	HMMER
			Protein-tyrosine phosphatase: N1264-E1495, N1553-E1786	HMMER_PFAM
			Fibronectin type III domain: P319-S401, P512-T594, P804-S889, P413-S500, P606-S698, P710-F793	HMMER_PFAM
			Immunoglobulin domain: G47-A109, G246-A300, T151-A209	HMMER_PFAM
			Cytosolic domain: K1174-T1796; Transmembrane domain: L1151-F1173; Non-cytosolic domain: M1-M1150	TMHMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Tyrosine specific protein phosphatases proteins BL00383: K1267-V1281, S1292-I1300, D1322-T1332, Q1397-P1409, V1435-G1445, R1473-F1488	BLIMPS_BLOCKS
			Tyrosine specific protein phosphatases signature and profiles: I1415-R1473	PROFILESCAN
			Tyrosine specific protein phosphatases signature and profiles: I1705-L1763	PROFILESCAN
			Synaptobrevin signature PR00219: M1150-A1169	BLIMPS_PRINTS
			Tissue factor signature PR00346: E1149-L1172	BLIMPS_PRINTS
			Protein tyrosine phosphatase signature PR00700: D1293-I1300, Y1309-E1329, R1393-E1410, P1432-I1450, V1754-A1769, M1479-I1489	BLIMPS_PRINTS
			PHOSPHATASE HYDROLASE RECEPTOR PRECURSOR SIGNAL PROTEIN TYROSINE PROTEIN TYROSINE GLYCOPROTEIN TRANSMEMBRANE PD000401: V965-W1152	BLAST_PRODOM
			HYDROLASE PHOSPHATASE PROTEIN PROTEIN TYROSINE PRECURSOR SIGNAL TYROSINE TRANSMEMBRANE GLYCOPROTEIN RECEPTOR PD000167: N1553-E1786	BLAST_PRODOM
			HYDROLASE PHOSPHATASE PROTEIN PROTEIN-TYROSINE TYROSINE PRECURSOR SIGNAL TRANSMEMBRANE GLYCOPROTEIN RECEPTOR PD000155: M1658-Y1787	BLAST_PRODOM
			PHOSPHATASE PROTEIN TYROSINE RECEPTOR PRECURSOR SIGNAL HYDROLASE TYPE- S PHOSPHOTYROSINE PD012459: K699-F807	BLAST_PRODOM
			Myb DNA-binding domain repeat signature 1: W1040-L1048	MOTIFS
			Tyrosine specific protein phosphatases active site: V1435-F1447, V1726-F1738	MOTIFS
			Leucine zipper pattern: L1151-L1172	MOTIFS
			Potential Phosphorylation Sites: S74 S75 S112 S119 S123 S196 S239 S455 S485 S819 S849 S926 S948 S981 S1041 S1116 S1182 S1183 S1259 S1390 S1474 S1537 S1587 S1608 S1647 T273 T314 T371 T568 T690 T691 T799 T871 T881 T984 T1037 T1042 T1088 T1178 T1284 T1319 T1339 T1367 T1378 T1526 T1542 T1683 T1759 T1762	MOTIFS
			Potential Glycosylation Sites: N117 N250 N295 N619 N855 N1585 N1620	MOTIFS
17	6818489CD1	438	signal_cleavage: M1-G26	SPSCAN
			Signal Peptide: M1-G23, M1-G26, M1-A31	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Purple acid phosphatase, N-terminal ig: S18-R124	HMMER_PFAM
			Purple acid phosphatase: N205-P384	HMMER_PFAM
			Aspartate carbamoyltransferase signature PR00101: M1-P5	BLIMPS_PRINTS
			Sodium/neurotransmitter symporter signature PR00176: H2-V21	BLIMPS_PRINTS
			Brain natriuretic peptide signature PR00712: S24-Q34	BLIMPS_PRINTS
			Laminin G domain protein PF00054: P5-L16	BLIMPS_PFAM
			ACID PHOSPHATASE PROTEIN PURPLE HYDROLASE PHOSPHO-MONOESTERASE PRECURSOR GLYCEROPHOSPHATASE SIGNAL IRON-III Zinc-II PD006329: M200-D418	BLAST_PRODOM
			PHOSPHATASE; II; PURPLE; IRON; DM08310|P80366|75-291: Y92-S292; |S51078|1-211: Y92-M284; |JC2545|292-446: S196-F248	BLAST_DOMO
			Potential Phosphorylation Sites: S118 S133 S292 T56 T97 T157 T300 Y397	MOTIFS
			Potential Glycosylation Sites: N211 N350 N404	MOTIFS
18	7509415CD1	635	Protein kinase domain: W34-F268	HMMER_PFAM
			SMRT_S_TKc domain: W34-F308	HMMER_PFAM
			Insulin-like growth factor BL00222: P432-V442	BLIMPS_BLOCKS
			Neutrophil cytosol factor P40 signature PR00497: L54-E65	BLIMPS_PRINTS
			Lysyl-tRNA synthetase signature PR00982:L37-E50	BLIMPS_PRINTS
			Diacylglycerol kinase catalytic domain PF00781: G18-K38	BLIMPS_PFAM
			PROTEIN GLUTATHIONE SYNT PD02939: K75-G86	BLIMPS_PRODOM
			SIMILAR TO CASEIN KINASES PD115501: F226-D316	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|P42169|74-330: I40-L278; |P48730|11-265: K38-Y286; |P35506|19-273: V36-Y286; |P40235|13-267: K38-Y286	BLAST_DOMO
			Protein kinases ATP-binding region signature: I40-K63	MOTIFS
			Potential Phosphorylation Sites: S291 S296 S365 S486 S535 S546 S550 S631 T131 T168 T186 T202 T282 T481	MOTIFS
			Potential Glycosylation Sites: N13 N185	MOTIFS
19	7506916CD1	444	signal_cleavage: M1-A24	SPSCAN
			Signal Peptide: M1-A24	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Protein kinase domain: F93-K331	HMMER_PFAM
			Tyrosine kinase catalytic domain signature PR00109: V170-Q183, Y206-L224, G252-F262, V276-H298	BLIMPS_PRINTS
			Octicosapeptide repeat protein PF00564: F93-E147, W157-L207, P352-L405	BLIMPS_PFAM
			PROTEIN KINASE DOMAIN DM00004|P54644|122-362: I95-Y295; DM00004|P28178|155-393: I95-L329; DM00004|P42818|136-373: I95-P339; DM00004|A53300|64-305: I95-K332	BLAST_DOMO
			Protein kinases ATP-binding region signature: I99-K122	MOTIFS
			Serine/Threonine protein kinases active-site signature: I212-L224	MOTIFS
			Potential Phosphorylation Sites: S3 S10 S159 S373 S375 T76 T115 T192 Y119	MOTIFS
			Potential Glycosylation Sites: N363	MOTIFS
20	7507104CD1	230	Protein phosphatase 2C proteins BL01032: R97-D110, D145-D157, S212-I221	BLIMPS_BLOCKS
			Potential Phosphorylation Sites: S112 S173 S212 T26 T47 T138 T160 T197 Y140	MOTIFS
			Potential Glycosylation Sites: N70	MOTIFS
21	7507105CD1	320	Protein phosphatase 2C: G73-R110, V194-A264	HMMER_PFAM
			Serine/threonine phosphatases, family 2C: V18-I311	HMMER_SMRT
			Protein phosphatase 2C proteins BL01032: R187-D200, D235-D247, S302-I311	BLIMPS_BLOCKS
			Potential Phosphorylation Sites: S55 S202 S263 S302 T116 T137 T228 T250 T287 Y230	MOTIFS
			Potential Glycosylation Sites: N160	MOTIFS
22	7507107CD1	184	Protein phosphatase 2C: S63-R128	HMMER_PFAM
			Protein phosphatase 2C proteins BL01032: L64-A81, G91-Y100, N109-E148	BLIMPS_BLOCKS
			Potential Phosphorylation Sites: S55 T134 T155	MOTIFS
			Potential Glycosylation Sites: N178	MOTIFS
23	7507109CD1	303	Protein phosphatase 2C: S63-R128, V212-Y273	HMMER_PFAM
			Protein phosphatase 2C proteins BL01032: L64-A81, R205-D218, D253-D265	BLIMPS_BLOCKS
			PROTEIN PHOSPHATASE 2C DM00377|P49596|1-295: S55-G143, R205-V272; DM00377|Q09173|1-296: V66-L141, E188-V272; DM00377|S62462|1-297: V66-L141, E188-V272	BLAST_DOMO
			Potential Phosphorylation Sites: S55 S220 T134 T155 T246 T268 Y248	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Glycosylation Sites: N178	MOTIFS
24	1833937CD1	53	Potential Phosphorylation Sites: S21 S28 S42 T50	MOTIFS
25	7502036CD1	495	Protein kinase domain: G13-V190	HMMER_PFAM
			Serine/threonine protein kinases, catalytic domain: M1-V190	HMMER_SMRT
			Protein kinases ATP-binding region proteins BL00107: H44-A74, P112-V127	BLIMPS_BLOCKS
			Protein kinases signatures and profile: D30-Q86	PROFILESCAN
			Tyrosine kinase catalytic domain signature PR00109: H44-L62, V113-E135, V159-A181	BLIMPS_PRINTS
			KINASE PROTEIN II CALCIUM/CALMODULIN-DEPENDENT TYPE SUBUNIT CHAIN TRANSFERASE SERINE/THREONINE PROTEIN CALMODULIN-BINDING PD004250: E407-Q495	BLAST_PRODOM
			KINASE PROTEIN II CALCIUM/CALMODULIN-DEPENDENT TYPE SUBUNIT CALMODULIN-BINDING CHAIN TRANSFERASE SERINE/THREONINE PROTEIN PD001779: V190-S247, T303-A324, R363-V406	BLAST_PRODOM
			CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE II ISOFORM GAMMA G PD063143: K257-A291	BLAST_PRODOM
			PROTEIN KINASE DOMAIN DM00004|JU0270|16-262: V11-A181; DM00004|P11798|15-261: V11-A181; DM00004|A44412|16-262: V11-A181	BLAST_DOMO
			KINASE; DEPENDENT; II; CALMODULIN; DM05068|P11798|263-426: Q183-R236, L251-V299, D304-K323, M362-A444	BLAST_DOMO
			Serine/Threonine protein kinases active-site signature: I50-L62	MOTIFS
			Potential Phosphorylation Sites: S2 S27 S334 S340 S464 T6 T12 T180 T290 T315 T316 T317 T395	MOTIFS
			Potential Glycosylation Sites: N231 N301 N314 N331	MOTIFS
26	7503248CD1	221	signal_cleavage: M41-G105	SPSCAN
			PROTEIN KINASE NUCLEAR HOMEODOMAIN-INTERACTING HOMEOBOX DNA-BINDING SERINE/THREONINE SERINE/THREONINE PROTEIN PD150854: M8-I154; PD150874: S178-I221	BLAST_PRODOM
			Potential Phosphorylation Sites: S27 S37 S121 S135 T119 T172	MOTIFS
			Potential Glycosylation Sites: N140 N157	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
27	7503968CD1	307	Protein kinase domain: V55-L173, W201-A293	HMMER_PFAM
			Serine/Threonine protein kinases, catalytic domain: Y28-A293	HMMER_SMRT
			Tyrosine kinase catalytic domain signature PR00109: Y147-L165, F197-L207, S215-E237, D266-I288	BLIMPS_PRINTS
			PROTEIN KINASE DOMAIN DM00004|S43968|28-311: Q33-A293, R271-I288; DM00004|A55480|28-320: Q33-A293, R271-V292; DM00004|P49186|28-320: Q33-A293, R271-V292; DM00004|P45984|28-320: Q33-A293, R271-V292	BLAST_DOMO
			Potential Phosphorylation Sites: S6 S20 S114 S212 S231 S244 S251 S283 T12 T183 T258 T269 T287	MOTIFS
			Potential Glycosylation Sites: N208	MOTIFS
28	7505931CD1	142	signal_cleavage: M1-S21	SPSCAN
			FGGY family of carbohydrate kinases, N-terminal: L12-P142	HMMER_PFAM
			FGGY family of carbohydrate kinases proteins BL00933: L12-L35, F47-P57	BLIMPS_BLOCKS
			GLYCEROL KINASE ATP: GLYCEROL 3-PHOSPHOTRANSFERASE GLYCEROKINASE GK METABOLISM TRANSFERASE POLYMORPHISM DISEASE PD014105:E106-P142	BLAST_PRODOM
			XYLULOKINASE DM02388|P32189|9-510: L9-V112, E106-M129; DM02388|P08859|2-493: A15-R117, P107-M129; DM02388|I64086|3-494: A15-R117, E115-M129; DM02388|P18157|1-492: A15-T120	BLAST_DOMO
			Potential Phosphorylation Sites: S4 S22 S63 T120 Y109	MOTIFS
29	7506912CD1	388	panK_eukar: pantothenate kinase: T49-E327	HMMER_TIGRFAM
			KINASE TRANSFERASE CDNA PANTOTHENATE FIS FOR T13D8.31 CODED COSMID ELEGANS PD018089: M31-D303, N289-L379	BLAST_PRODOM
			Potential Phosphorylation Sites: S4 S62 S115 S298 S300 S345 T49 T182 T224 T241 T245 T259	MOTIFS
			Potential Glycosylation Sites: N87 N118	MOTIFS
30	7506913CD1	314	KINASE TRANSFERASE CDNA PANTOTHENATE FIS FOR T13D8.31 CODED COSMID ELEGANS PD018089: M16-G214, N215-L305	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Phosphorylation Sites: S47 S100 S224 S226 S271 T10 T34 T167 T209	MOTIFS
			Potential Glycosylation Sites: N72 N103	MOTIFS
31	7507029CD1	166	Dual specificity phosphatase, catalytic domain: R22-G152	HMMER_PFAM
			Tyrosine specific protein phosphatases proteins BL00383: V97-A107	BLIMPS_BLOCKS
			Tyrosine specific protein phosphatases signature and profiles: L74-P134	PROFILESCAN
			Protein tyrosine phosphatase signature PR00700: V94-G112	BLIMPS_PRINTS
			VH1-TYPE DUAL SPECIFICITY PHOSPHATASE DM03823|I38890|29-320: Q78-Q150	BLAST_DOMO
			Potential Phosphorylation Sites: S155 T27 T34 T86 T148	MOTIFS
			Potential Glycosylation Sites: N164	MOTIFS
32	7507063CD1	173	signal_cleavage: M49-D112	SPSCAN
			Rhodanese proteins BL00380: L24-S35, S35-I45, A80-W92	BLIMPS_BLOCKS
			PHOSPHATASE DUAL SPECIFICITY PROTEIN HYDROLASE KINASE MAP PYST1 DUSP6ALT MITOGEN-ACTIVATED PD021468: G47-Q121	BLAST_PRODOM
			Potential Phosphorylation Sites: S35 T156	MOTIFS
33	7504755CD1	486	Signal Peptide: M1-C19, M1-E26	HMMER
			Cytosolic domain: G449-V486; Transmembrane domain: V426-L448; Non-cytosolic domain: M1-L425	TMHMMER
			TYROSINE KINASE LEUKOCYTE RECEPTOR PRECURSOR TRANSFERASE TYROSINE PROTEIN TRANSMEMBRANE ATP-BINDING PHOSPHORYLATION PD023963: M1-C300; PD022408: D335-L448	BLAST_PRODOM
			RECEPTOR TYROSINE KINASE CLASS II DM04844|P29376|1-211: M1-G212; DM04844|B48266|1-221: M1-G208; DM04844|P08923|1-208: M1-G212, G265-G276, G308-G320, A324-G330; DM04843|P29376|314-509: G206-T215, G314-L448	BLAST_DOMO
			Potential Phosphorylation Sites: S40 S51 S76 S201 S289 S337 T84 T278 T414	MOTIFS
			Potential Glycosylation Sites: N257 N380 N412	MOTIFS
34	7509265CD1	256	Protein kinase domain: Y4-R180	HMMER_PFAM
			Protein kinases ATP-binding region proteins BL00107: F130-G160	BLIMPS_BLOCKS
			Protein kinases signatures and profile: A82-D168	PROFILESCAN
			Tyrosine kinase catalytic domain signature PR00109: F130-I148	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			PROTEIN KINASE DOMAIN DM00004|P29620|21-289: I10-R180; DM00004|P23437|6-286: R9-T179; DM00004|P43450|6-276: R9-T179; DM00004|I48288|6-277: R9-T179	BLAST_DOMO
			Serine/Threonine protein kinases active-site signature: I136-I148	MOTIFS
			Potential Phosphorylation Sites: S184 S250 T188	MOTIFS
35	7509371CD1	458	Protein kinase domain: L44-C282	HMMER_PFAM
			Protein kinases ATP-binding region proteins BL00107: H142-F172, K207-S222	BLIMPS_BLOCKS
			Receptor tyrosine kinase class II proteins BL00239: E75-S122, S179-E228, L232-I276	BLIMPS_BLOCKS
			Protein kinases signatures and profile: I129-S182	PROFILESCAN
			Tyrosine kinase catalytic domain signature PR00109: M104-Q117, H142-L160, S208-A230, Y251-L273	BLIMPS_PRINTS
			PROTEIN KINASE DOMAIN DM00004|P51954|6-248: D50-P271; DM00004|P51957|8-251: V42-P271; DM00004|Q08942|22-269: D50-P271; DM00004|P51955|10-261: R47-P271	BLAST_DOMO
			Serine/Threonine protein kinases active-site signature: I148-L160	MOTIFS
			Potential Phosphorylation Sites: S122 S179 S222 S248 S295 S316 S396 S419 T27 T65 T111 T346 T411	MOTIFS
36	7509389CD1	362	Protein kinase domain: S64-C160	HMMER_PFAM
			Protein kinases ATP-binding region proteins BL00107: K85-S100	BLIMPS_BLOCKS
			Tyrosine kinase catalytic domain signature PR00109: S86-A108, Y129-L151	BLIMPS_PRINTS
			PROTEIN KINASE DOMAIN DM00004|P51954|6-248: L55-P149; DM00004|P51957|8-251: L55-P149; DM00004|Q08942|22-269: G67-P149; DM00004|P11837|13-285: V66-P149	BLAST_DOMO
			Potential Phosphorylation Sites: S10 S24 S28 S43 S61 S100 S126 S173 S300 S323 T47 T315	MOTIFS
37	7507005CD1	292	signal_cleavage: M1-C34	SPSCAN
			PUTATIVE PHOSPHATASE PD123939: C33-K290	BLAST_PRODOM
			Potential Phosphorylation Sites: S83 S157 S175 S267 T61 T87 T277	MOTIFS
38	7509142CD1	226	Tyrosine specific protein phosphatases proteins BL00383: K186-Y200	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			PROTEIN TYROSINE PHOSPHATASE, NONRECEPTOR TYPE 20 EC 3.1.3.48 PHOSPHOTYROSINE PHOSPHATASE PTPASE HYDROLASE PD097276: M1-N179	BLAST_PRODOM
			Potential Phosphorylation Sites: S3 S20 S34 S51 S57 S76 S95 S96 S120 S182 T39 T84	MOTIFS
			Potential Glycosylation Sites: N18	MOTIFS
39	7509157CD1	261	Protein-tyrosine phosphatase: N24-E252	HMMER_PFAM
			Tyrosine specific protein phosphatases proteins BL00383: K27-V41, K47-I55, D79-S89, Q156-P168, V192-G202, R230-F245	BLIMPS_BLOCKS
			Tyrosine specific protein phosphatases signature and profiles: A181-M229	PROFILESCAN
			Protein tyrosine phosphatase signature PR00700: D48-I55, Y66-E86, H152-A169, P189-V207, F220-G235, M236-C246	BLIMPS_PRINTS
			HYDROLASE PHOSPHATASE PROTEIN PROTEIN TYROSINE PRECURSOR SIGNAL TYROSINE TRANSMEMBRANE GLYCOPROTEIN RECEPTOR PD000167: N24-V207	BLAST_PRODOM
			HYDROLASE PHOSPHATASE PROTEIN PROTEIN TYROSINE TYROSINE PRECURSOR SIGNAL TRANSMEMBRANE GLYCOPROTEIN RECEPTOR PD000155: L127-S234, T167-K256	BLAST_PRODOM
			PROTEIN-TYROSINE-PHOSPHATASE DM00089|A54971|2204-2475: E8-L257; DM00089|P34138|13-350: E8-D48, D48-L254; DM00089|P26045|632-904: P22-V253; DM00089|P23468|1623-1906: M5-L254	BLAST_DOMO
			Tyrosine specific protein phosphatases active site: V192-F204	MOTIFS
			Potential Phosphorylation Sites: S23 S38 S75 S113 S153 T76 Y36	MOTIFS
			Potential Glycosylation Sites: N51 N87 N217	MOTIFS
40	7509246CD1	173	signal_cleavage: M49-D112	SPSCAN
			Rhodanese proteins BL00380: L24-S35, S35-I45, A80-W92	BLIMPS_BLOCKS
			PHOSPHATASE DUAL SPECIFICITY PROTEIN HYDROLASE KINASE MAP PYST1 DUSP6ALT MITOGENACTIVATED PD021468: G47-Q121	BLAST_PRODOM
			Potential Phosphorylation Sites: S35 T156	MOTIFS
41	7509380CD1	412	Protein-tyrosine phosphatase: N183-E406	HMMER_PFAM
			Protein tyrosine phosphatase, catalytic domain: D158-K409	HMMER_SMRT

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Protein tyrosine phosphatase, catalytic motif: S312-X413	HMMER_SMRT
			Tyrosine specific protein phosphatases proteins BL00383: K186-V200, K206-I214, D238-S248, Q315-P327, V351-G361, R389-V404	BLIMPS_BLOCKS
			Tyrosine specific protein phosphatases signature and profiles: A340-M388	PROFILESCAN
			Protein tyrosine phosphatase signature PR00700: D207-I214, Y225-E245, H311-A328, P348-V366, F379-G394	BLIMPS_PRINTS
			PROTEIN TYROSINE PHOSPHATASE, NONRECEPTOR TYPE 20 EC 3.1.3.48 PHOSPHOTYROSINE PHOSPHATASE PTPASE HYDROLASE PD097276: M1-N179	BLAST_PRODOM
			HYDROLASE PHOSPHATASE PROTEIN PROTEIN TYROSINE PRECURSOR SIGNAL TYROSINE TRANSMEMBRANE GLYCOPROTEIN RECEPTOR PD000167: N183-V366	BLAST_PRODOM
			HYDROLASE PHOSPHATASE PROTEIN PROTEIN TYROSINE TYROSINE PRECURSOR SIGNAL TRANSMEMBRANE GLYCOPROTEIN RECEPTOR PD000155: L286-E400	BLAST_PRODOM
			PROTEIN-TYROSINE-PHOSPHATASE DM00089|A54971|2204-2475: E167-E400; DM00089|P26045|632-904: I159-T398; DM00089|P34138|13-350: E167-E400; DM00089|P23468|1623-1906: M164-E400	BLAST_DOMO
			Tyrosine specific protein phosphatases active site: V351-F363	MOTIFS
			Potential Phosphorylation Sites: S3 S20 S34 S51 S57 S76 S95 S96 S120 S182 S197 S234 S272 S312 S403 T39 T84 T115 T235 Y195	MOTIFS
			Potential Glycosylation Sites: N18 N210 N246 N376	MOTIFS
42	7509382CD1	197	PROTEIN TYROSINE PHOSPHATASE, NONRECEPTOR TYPE 20 EC 3.1.3.48 PHOSPHOTYROSINE PHOSPHATASE PTPASE HYDROLASE PD097276: M1-N179	BLAST_PRODOM
			Potential Phosphorylation Sites: S3 S20 S34 S51 S57 S76 S95 S96 S120 S182 T39 T84 T115 Y195	MOTIFS
			Potential Glycosylation Sites: N18	MOTIFS

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
43/ 7503679CB1/ 3853	1-726, 14-3853, 414-1301, 457-1306, 474-1306, 478-1306, 485-1299, 485-1301, 485-1302, 487-1306, 496-1306, 502-1306, 508-1200, 508-1306, 513-1306, 514-1302, 515-1306, 524-1306, 525-1306, 526-1298, 536-1302, 547-1310, 556-1298, 579-1306, 584-1306, 614-1312, 627-1306, 658-1306, 684-1306, 901-1306, 933-1302, 933-1306, 1086-1301, 1086-1306, 1256-1883, 1256-1907, 1256-1914, 1256-1944, 1256-1972, 1258-1989, 1317-1987, 1349-1987, 1369-1584, 1369-1592, 1369-1593, 1369-1594, 1369-1926, 1369-1984, 1371-1998, 1379-1984, 1383-1982, 1537-1987, 1576-1987, 1983-2510, 1983-2521, 2031-2521, 2032-2521, 2207-2913, 2246-2916, 2453-2521, 2611-2916, 2637-3340, 2658-3146, 2662-3293, 2707-3511, 2726-3344, 2728-3451, 2732-3182, 2755-2914, 2774-3389, 2783-3347, 2787-3395, 2824-3435, 2840-3360, 2847-3438, 2847-3522, 2849-3598, 2854-3433, 2864-3317, 2880-3477, 2880-3606, 2901-
	3645, 2912-3156, 2917-3185, 2917-3308, 2922-3626, 2936-3551, 2962-3544, 2962-3646, 2962-3647, 2978-3425, 2984-3478, 2994-3704, 3004-3263, 3013-3540, 3051-3853, 3058-3797, 3065-3663, 3068-3853, 3070-3665, 3075-3853, 3108-3735, 3112-3399, 3120-3661, 3127-3853, 3136-3783, 3142-3847, 3148-3376, 3148-3775, 3165-3374, 3171-3582, 3172-3788, 3177-3582, 3187-3853, 3208-3853, 3212-3423, 3226-3335, 3228-3694, 3250-3853, 3315-3550, 3413-3667
44/ 7503681CB1/ 4000	1-755, 1-4000, 243-1130, 286-1135, 303-1135, 307-1135, 314-1128, 314-1130, 314-1131, 316-1135, 325-1135, 331-1135, 337-1029, 337-1135, 342-1135, 343-1131, 344-1135, 353-1135, 354-1135, 355-1127, 365-1131, 376-1139, 385-1127, 408-1135, 413-1135, 443-1141, 456-1135, 487-1135, 513-1135, 730-1135, 762-1131, 762-1135, 915-1130, 915-1135, 1085-1712, 1085-1736, 1085-1743, 1085-1773, 1085-1801, 1087-1818, 1146-1816, 1178-1816, 1198-1413, 1198-1421, 1198-1422, 1198-1423, 1198-1755, 1198-1813, 1200-1827, 1208-1813, 1212-1811, 1366-1816, 1405-1816, 1814-2362, 1974-2668, 1987-2668, 2000-2643, 2034-2668, 2058-2668, 2075-2668, 2080-2665, 2085-2668, 2087-2668, 2088-2668, 2095-2668, 2097-2668, 2178-2668, 2206-2668, 2241-3060, 2280-2855, 2281-2665, 2312-2668, 2393-3063, 2600-2668, 2784-3487, 2805-3293, 2809-3440, 2854-3658, 2873-3491, 2875-3598, 2879-3329, 2902-3061, 2921-3536, 2930-3494, 2934-3542, 2971-3582, 2987-3507, 2994-3585, 2994-3669, 2996-3745,
	3001-3580, 3011-3464, 3027-3624, 3027-3753, 3048-3792, 3059-3303, 3064-3332, 3064-3455, 3069-3773, 3083-3698, 3109-3691, 3109-3793, 3109-3794, 3125-3572, 3131-3625, 3141-3851, 3151-3410, 3160-3687, 3198-4000, 3205-3944, 3212-3810, 3215-4000, 3217-3812, 3222-4000, 3255-3882, 3259-3546, 3267-3808, 3274-4000, 3283-3930, 3289-3994, 3295-3523, 3295-3922, 3312-3521, 3318-3729, 3319-3935, 3324-3729, 3334-4000, 3355-4000, 3359-3570, 3373-3482, 3375-3841, 3397-4000, 3462-3697, 3560-3814

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
45/ 7505819CB1/ 6408	1-241, 3-545, 5-598, 7-530, 9-444, 17-598, 30-241, 31-306, 31-313, 31-364, 31-537, 31-542, 31-593, 31-594, 31-595, 31-596, 31-598, 31-6408, 34-608, 35-598, 38-561, 38-598, 41-444, 41-598, 48-598, 50-515, 78-390, 99-596, 104-386, 105-598, 109-598, 125-550, 175-592, 177-598, 178-592, 187-598, 215-539, 217-596, 233-411, 246-509, 259-463, 260-485, 268-427, 285-598, 297-486, 334-598, 337-608, 339-592, 366-594, 373-591, 466-608, 493-591, 522-598, 584-881, 584-1032, 598-1154, 628-1308, 724-934, 756-1558, 829-1373, 829-1381, 855-1526, 1079-1655, 1229-1488, 1279-1822, 1386-1495, 1445-1613, 1446-1967, 1461-1711, 1479-2018, 1480-2121, 1536-2111, 1599-2201, 1615-2189, 1690-2243, 1729-1968, 1740-2371, 1756-1903, 1803-2304, 1887-2374, 1887-2511, 1887-2516, 1917-2525, 2051-2484, 2069-2289, 2102-2601, 2125-2601, 2188-2601, 2224-2601, 2241-2480, 2406-2927, 2413-2679, 2416-2603, 2423-2604, 2609-3218, 2626-2894, 2697-3408, 2771-3285, 2840-3101, 2914-3178, 3018-3609, 3121-3429,
	3271-3732, 3278-3670, 3385-4210, 3458-3673, 3473-3697, 3489-3752, 3490-4205, 3503-3620, 3503-4015, 3503-4016, 3599-4008, 3644-4082, 3644-4240, 3656-3936, 3657-4263, 3678-3997, 3689-4542, 3700-3965, 3746-4311, 3746-4467, 3748-4125, 3751-3992, 3898-3999, 3921-4621, 3966-4278, 4126-4390, 4133-4366, 4144-4746, 4154-4658, 4156-4591, 4166-4804, 4181-4773, 4217-4484, 4225-4819, 4228-4794, 4291-4819, 4349-4862, 4384-4710, 4414-4542, 4414-4827, 4414-4933, 4467-4784, 4474-4817, 4552-4828, 4595-4795, 4595-4810, 4595-4947, 4607-4813, 4615-4765, 4627-4896, 4644-4807, 4651-5196, 4684-4826, 4709-5384, 4831-5106, 4832-5036, 4836-4954, 4844-5103, 4844-5372, 4845-5101, 4852-5077, 4865-5306, 4866-5369, 4887-5481, 4917-5479, 4934-5368, 4948-5490, 4968-5296, 4968-5396, 4986-5364, 4999-5252, 5003-5178, 5007-5482, 5030-5322, 5030-5539, 5051-5318, 5061-5242, 5075-5336, 5076-5535, 5077-5329, 5077-5365, 5109-5320, 5109-5347, 5132-5346, 5150-5754,
	5179-5398, 5179-5405, 5181-5411, 5183-5456, 5188-5346, 5204-5281, 5206-5281, 5244-5770, 5322-5549, 5322-5591, 5327-5579, 5361-5836, 5431-5962, 5488-6088, 5591-6353, 5596-6352, 5609-6138, 5629-5898, 5630-6168, 5634-5890, 5648-5904, 5651-5925, 5652-5962, 5653-5952, 5659-6220, 5664-5940, 5664-6208, 5695-5950, 5695-5951, 5698-5957, 5712-5928, 5713-6283, 5721-5981, 5722-6332, 5736-5995, 5737-6398, 5738-5987, 5741-5950, 5741-5963, 5756-6034, 5757-6366, 5764-6006, 5766-6009, 5793-6148, 5793-6342, 5795-6005, 5796-6032, 5797-5997, 5798-6360, 5800-6407, 5806-6046, 5806-6047, 5806-6153, 5809-6136, 5828-6204, 5833-6147, 5834-6053, 5837-6319, 5839-6102, 5841-6127, 5844-6283, 5856-6377, 5860-6160, 5865-6148, 5871-6225, 5873-6145, 5877-6108, 5883-6258, 5907-6380, 5915-6405, 5925-6394, 5926-6394, 5934-6398, 5939-6397, 5942-6408, 5944-6408, 5948-6408, 5953-6080, 5955-6394, 5956-6395, 5960-6394, 5962-6394, 5974-6408, 5979-6241, 5979-6397,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	5981-6394, 5982-6394, 5982-6400, 5985-6400, 5986-6394, 5987-6394, 5989-6408, 5991-6398, 5993-6400, 5999-6408, 6001-6399, 6009-6331, 6011-6394, 6018-6394, 6018-6400, 6027-6296, 6032-6394, 6037-6394, 6038-6394, 6039-6399, 6054-6289, 6054-6354, 6054-6394, 6059-6384, 6063-6392, 6068-6394, 6069-6333, 6072-6400, 6072-6401, 6085-6408, 6088-6328, 6089-6402, 6095-6394, 6098-6394, 6110-6408, 6113-6360, 6118-6395, 6119-6398, 6122-6344, 6127-6283, 6147-6343, 6155-6386, 6157-6399, 6161-6394, 6162-6394, 6167-6298, 6180-6391, 6183-6408, 6186-6394, 6188-6405, 6196-6265, 6197-6408, 6202-6408, 6205-6376, 6210-6399, 6219-6394, 6226-6394, 6243-6400, 6247-6394, 6262-6394, 6278-6394, 6330-6394
46/ 7505083CB1/ 1072	1-129, 14-1072, 27-129, 27-734, 27-929, 27-933, 29-118, 133-1044, 141-1043, 175-1043, 185-1043, 219-1043, 255-1043, 286-1043, 290-1043, 332-1043, 369-1039, 477-1043, 771-1043
47/ 7505866CB1/ 4822	1-549, 1-643, 1-665, 1-724, 1-776, 1-778, 1-870, 1-912, 1-4453, 5-685, 57-897, 349-852, 356-574, 549-1085, 658-920, 662-1492, 770-1346, 849-1464, 980-1169, 987-1284, 988-1499, 988-1513, 988-1562, 988-1587, 988-1710, 988-1822, 993-1715, 1017-1668, 1025-1520, 1036-1585, 1046-1700, 1049-1585, 1119-1263, 1166-1690, 1205-1352, 1240-1822, 1260-1687, 1355-1977, 1399-1693, 1481-1960, 1555-1878, 1582-2207, 1610-2222, 1716-2008, 1718-2007, 1734-2197, 1767-2493, 1786-2604, 1833-2454, 1863-2281, 1896-2519, 1971-2735, 2014-2282, 2020-2596, 2029-2624, 2160-2430, 2263-2706, 2292-2544, 2410-2646, 2410-2652, 2410-2693, 2410-2943, 2412-2729, 2419-2642, 2469-2743, 2494-2956, 2502-2713, 2502-2748, 2503-2886, 2554-3081, 2660-2946, 2676-2886, 2676-2891, 2679-2918, 2703-2918, 2722-3348, 2722-3353, 2813-3100, 2863-3022, 2878-3114, 2878-3130, 2896-3142, 2966-3213, 3137-3919, 3140-3695, 3141-3919, 3155-3920, 3159-3919, 3167-3919, 3173-3919, 3181-3919, 3197-3617,
	3212-3676, 3213-3802, 3217-3919, 3223-3444, 3244-3489, 3245-3966, 3246-3798, 3253-3491, 3255-3546, 3278-3987, 3294-3559, 3296-3439, 3311-3819, 3316-3591, 3323-3551, 3340-3580, 3340-3608, 3340-3616, 3357-3594, 3360-3948, 3360-3956, 3372-3721, 3389-3772, 3400-3675, 3400-3680, 3402-3665, 3402-3995, 3407-3699, 3416-3740, 3437-3710, 3437-3747, 3438-3828, 3440-4119, 3465-3772, 3482-3697, 3487-3892, 3489-3740, 3494-3658, 3499-3747, 3509-3768, 3532-3802, 3537-3763, 3570-3770, 3572-3711, 3572-3840, 3573-4223, 3587-3863, 3587-4099, 3590-4248, 3608-4283, 3609-4231, 3612-3910, 3618-4060, 3626-4112, 3647-3903, 3651-3770, 3652-3917, 3663-3963, 3669-4228, 3679-3873, 3679-3963, 3688-4008, 3708-4020, 3725-3937, 3727-3957, 3736-4105, 3738-4231, 3739-4204, 3752-4388, 3756-4251, 3759-4143, 3760-4250, 3770-4411, 3777-4004, 3779-4071, 3779-4440, 3781-4391, 3784-4197, 3786-4253, 3790-4252, 3793-4213, 3798-4024, 3803-4043, 3807-3919, 3807-4212,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	3807-4231, 3807-4250, 3832-4253, 3846-4453, 3850-4105, 3851-4100, 3852-4118, 3853-4251, 3873-4208, 3877-4250, 3878-4220, 3891-4211, 3897-4248, 3900-4453, 3913-4351, 3932-4248, 3937-4219, 3940-4248, 3943-4251, 3954-4423, 3964-4230, 3975-4470, 3978-4253, 3979-4460, 3982-4479, 3987-4253, 3988-4466, 3994-4461, 4009-4231, 4009-4455, 4009-4480, 4010-4250, 4012-4468, 4024-4301, 4033-4250, 4036-4460, 4036-4464, 4047-4251, 4058-4247, 4060-4435, 4072-4459, 4075-4355, 4075-4471, 4080-4491, 4098-4462, 4123-4253, 4124-4330, 4131-4463, 4140-4459, 4164-4478, 4172-4742, 4208-4246, 4208-4249, 4319-4822, 4338-4572, 4342-4461, 4348-4821, 4369-4479, 4375-4455
48/ 7503214CB1/ 4368	1-4368, 919-1331, 930-1678, 930-1747, 930-1801, 930-1836, 931-1423, 931-1655, 931-1676, 931-1713, 931-1751, 931-1757, 931-1882, 934-1660, 935-1562, 948-1858, 955-1858, 1039-1344, 1053-1889, 1100-1698, 1100-1722, 1100-1742, 1100-1837, 1103-1771, 1136-1829, 1222-1593, 1330-2046, 1338-2050, 1343-2045, 1351-2048, 1365-2046, 1376-1583, 1380-2041, 1382-2049, 1396-1639, 1399-2052, 1413-2050, 1421-2050, 1425-2049, 1428-2050, 1430-2046, 1430-2050, 1433-2050, 1437-2050, 1444-2047, 1444-2050, 1446-2049, 1449-2050, 1454-2050, 1462-2050, 1466-2050, 1468-1693, 1475-1741, 1487-2052, 1591-2050, 1658-2050, 1698-1862, 1873-2038, 1888-2296, 1888-2401, 1888-2514, 1892-2436, 1895-2300, 1899-2400, 1911-2401, 1919-2401, 1922-2401, 1925-2158, 1925-2161, 1925-2459, 1927-2401, 1948-2649, 1951-2642, 1969-2641, 2018-2649, 2031-2401, 2068-2290, 2068-2292, 2159-2401, 2335-2887, 2667-2987, 2681-3653, 2736-3496, 2751-3467, 2786-3486, 2786-3487, 2788-3482, 2789-
	3462, 2810-3648, 2822-3659, 2824-3409, 2826-3602, 2844-3351, 2894-3486, 2958-3116, 2972-3620, 3028-3721, 3043-3616, 3049-3891, 3071-3785, 3078-3567, 3080-3606, 3089-3401, 3140-3662, 3158-3736, 3164-3690, 3171-3784, 3190-3878, 3201-3510, 3207-3830, 3209-4051, 3218-3779, 3221-3933, 3238-4004, 3287-3949, 3317-3887, 3325-4232, 3358-3684, 3362-4088, 3367-4036, 3369-4024, 3379-3665, 3386-4107, 3402-3696, 3413-4088, 3438-4104, 3446-4086, 3455-3974, 3469-4153, 3471-4113, 3497-3976, 3514-4095, 3520-4138, 3538-4112, 3559-4219, 3574-3864, 3577-4104, 3609-4142, 3611-4073, 3624-4132, 3642-4319, 3687-4368, 3709-4368, 3741-4349, 3802-4128, 3811-4368, 3839-4368, 4080-4353
49/ 7495312CB1/ 2024	1-319, 234-721, 237-678, 269-720, 274-785, 287-949, 308-809, 309-948, 321-910, 337-783, 420-883, 525-1189, 565-1121, 582-1212, 586-1028, 592-1115, 595-1194, 618-924, 652-1151, 685-1261, 764-1348, 771-1356, 838-1485, 861-1375, 881-1509, 896-1520, 914-1366, 916-1500, 972-1578, 975-1545, 984-1509, 1072-1699, 1094-1644, 1123-1776, 1142-1461, 1142-1792, 1145-1664, 1156-1686, 1171-1700, 1173-1443, 1173-1528, 1202-1731, 1208-1706, 1264-1817, 1287-1814, 1393-1974, 1397-2018, 1460-1797, 1472-1951, 1477-2000, 1511-1936, 1532-1817, 1632-1817, 1634-2024, 1706-1904

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
50/ 7506732CB1/ 715	1-236, 1-247, 1-715, 34-235, 38-276, 42-358, 51-358, 54-293, 66-340, 73-343, 82-154, 82-358, 88-337, 96-322, 108-343, 109-355, 136-715, 149-358, 229-358, 353-465, 353-468, 353-525, 353-537, 353-568, 353-570, 353-571, 353-577, 353-586, 353-593, 353-595, 353-596, 353-598, 353-602, 353-605, 353-607, 353-611, 353-612, 353-615, 353-618, 353-620, 353-628, 353-715, 355-464, 356-505, 356-594, 356-596, 356-623, 358-622, 359-554, 360-598, 361-564, 364-590, 365-616, 365-642, 366-588, 369-559, 370-644, 371-639, 374-509, 374-555, 376-635, 376-648, 379-650, 383-673, 388-594, 388-599, 390-675, 399-617, 404-618, 404-669, 406-661, 413-653, 415-715, 426-686, 426-689, 432-696, 432-698, 435-596, 435-689, 438-663, 441-682, 441-697, 441-704, 441-714, 443-664, 444-695, 456-657, 458-644, 459-684, 460-648, 460-714, 460-715, 465-684, 474-715, 477-673, 477-704, 480-625, 480-701, 485-708, 489-715, 491-606, 497-596, 546-712, 549-632, 557-715, 560-706, 575-685
51/ 7506736CB1/ 591	1-66, 1-236, 1-247, 1-591, 15-66, 34-235, 38-276, 38-458, 38-460, 42-396, 42-466, 48-385, 51-360, 54-293, 58-429, 58-456, 58-466, 66-340, 73-343, 74-466, 82-154, 82-358, 82-396, 82-466, 88-337, 96-322, 106-398, 108-343, 109-387, 123-397, 124-392, 129-436, 136-384, 136-591, 137-343, 149-427, 183-438, 205-397, 227-466, 229-387, 229-407, 229-445, 229-466, 231-438, 231-466, 232-442, 232-466, 234-466, 235-413, 236-466, 249-446, 260-380, 277-462, 277-466, 344-384, 405-425, 405-429, 464-561, 464-582
52/ 7507121CB1/ 7424	1-493, 1-528, 1-556, 1-576, 1-580, 1-640, 1-642, 1-691, 1-749, 1-761, 6-667, 9-718, 24-696, 42-521, 42-550, 42-765, 44-501, 44-549, 44-573, 44-650, 44-665, 44-770, 44-891, 45-506, 172-798, 183-876, 271-883, 286-926, 293-626, 294-969, 305-453, 308-626, 313-973, 320-894, 332-462, 333-570, 333-587, 340-610, 342-1047, 348-926, 359-1052, 366-570, 367-570, 378-668, 388-819, 391-618, 396-1008, 408-1067, 472-1276, 487-1112, 495-1121, 552-1217, 561-1212, 601-1168, 620-1269, 624-1068, 631-1230, 635-1301, 649-1191, 661-1284, 661-1326, 666-1321, 670-1256, 689-1209, 690-1366, 694-1351, 697-1277, 703-1239, 705-1267, 712-1227, 719-1282, 723-1383, 746-1273, 747-1076, 751-1150, 754-1331, 763-1384, 794-1262, 797-1507, 803-1324, 808-1097, 818-1409, 830-1137, 837-1376, 851-1373, 854-1295, 876-1261, 880-1571, 883-1599, 887-1517, 888-1589, 893-1121, 912-1261, 920-1485, 920-1488, 964-1372, 966-1543, 973-1313, 978-1619, 989-1571, 1022-1262, 1024-1566, 1028-1551, 1030-1465, 1051-1262, 1052-1327, 1052-1334,
	1052-1385, 1052-1430, 1052-1539, 1052-1558, 1052-1563, 1052-1614, 1052-1615, 1052-1616, 1052-1617, 1052-1619, 1052-1629, 1052-7424, 1054-1619, 1055-1629, 1056-1619, 1059-1582, 1062-1586, 1071-1536, 1099-1411, 1120-1617, 1126-1619, 1130-1619, 1146-1571, 1605-1902, 1605-2053, 1618-2175, 1649-2329, 1745-1955, 1777-2579, 1850-2394, 1850-2402, 1876-2547, 2100-2676, 2250-2509, 2300-2843, 2407-2516, 2466-2634, 2467-2988, 2482-2732, 2500-3039, 2501-3142, 2557-3132, 2620-3222, 2636-3210, 2711-3264, 2750-2989, 2764-2786, 2824-3325, 2908-3395, 2908-3532, 2908-3537, 2938-3546, 3072-3505, 3090-3310, 3123-3622, 3146-3622, 3209-3622, 3245-3622, 3262-3501, 3427-3948, 3434-3700, 3437-3624, 3444-3625, 3630-4239, 3647-3915, 3718-4429, 3792-4306, 3861-4122, 3935-4199, 4039-4630, 4142-4450, 4292-4753, 4299-4691, 4406-5231, 4479-4694, 4494-4718, 4510-4773, 4511-5226, 4524-4641, 4524-5036, 4524-5037, 4620-5029, 4665-5103, 4665-5261, 4677-4957,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	4678-5284, 4699-5018, 4710-5563, 4721-4986, 4767-5332, 4767-5488, 4769-5146, 4772-5013, 4919-5020, 4942-5642, 4987-5299, 5147-5411, 5154-5387, 5165-5767, 5175-5679, 5177-5612, 5187-5825, 5202-5794, 5238-5505, 5246-5840, 5249-5815, 5312-5840, 5370-5883, 5405-5731, 5435-5563, 5435-5848, 5435-5954, 5488-5805, 5495-5838, 5573-5849, 5616-5816, 5616-5831, 5616-5968, 5628-5834, 5636-5786, 5648-5917, 5665-5828, 5672-6217, 5705-5847, 5730-6405, 5852-6127, 5853-6057, 5857-5975, 5865-6124, 5865-6393, 5866-6122, 5873-6098, 5886-6327, 5887-6390, 5908-6502, 5938-6500, 5955-6389, 5969-6511, 5989-6317, 5989-6417, 6007-6385, 6020-6273, 6024-6199, 6028-6503, 6051-6343, 6051-6560, 6072-6339, 6082-6263, 6096-6357, 6097-6556, 6098-6350, 6098-6386, 6200-6419, 6200-6426, 6202-6432, 6204-6477, 6225-6302, 6612-7374, 6617-7373, 6630-7159, 6650-6919, 6651-7189, 6655-6911, 6669-6925, 6672-6946, 6680-7241, 6685-7229, 6734-7304, 6743-7353, 6758-7419,
	6778-7387, 6814-7169, 6814-7363, 6819-7381, 6821-7424, 6827-7174, 6830-7157, 6849-7225, 6858-7340, 6862-7148, 6865-7304, 6877-7398, 6881-7181, 6886-7169, 6892-7246, 6894-7166, 6904-7279, 6928-7401, 6936-7424, 6946-7415, 6947-7415, 6955-7419, 6960-7418, 6963-7424, 6965-7424, 6969-7424, 6976-7415, 6977-7416, 6981-7415, 6983-7415, 6995-7424, 7000-7262, 7000-7418, 7002-7415, 7003-7415, 7003-7421, 7006-7421, 7007-7415, 7008-7415, 7010-7424, 7012-7419, 7014-7421, 7022-7420, 7030-7352, 7032-7415, 7039-7415, 7039-7421, 7048-7317, 7053-7415, 7058-7415, 7059-7415, 7060-7420, 7075-7310, 7075-7375, 7075-7415, 7080-7405, 7084-7413, 7089-7415, 7090-7354, 7093-7421, 7093-7422, 7106-7424, 7109-7349, 7110-7423, 7116-7415, 7119-7415, 7131-7424, 7134-7381, 7139-7416, 7140-7419, 7143-7365, 7148-7304, 7168-7364, 7176-7407, 7178-7420, 7182-7415, 7183-7415, 7188-7299, 7201-7412, 7204-7424, 7207-7415, 7209-7419, 7217-7286, 7218-7424, 7223-7424, 7226-7397, 7231-7420, 7240-7415, 7247-7415, 7264-7421, 7268-7415, 7283-7415, 7299-7415, 7351-7415
53/ 90086258CB1/ 1303	1-709, 1-1303, 401-1299
54/ 1967990CB1/ 2111	1-630, 30-320, 32-770, 45-337, 66-660, 137-523, 138-523, 139-686, 156-433, 206-464, 206-684, 206-727, 206-754, 206-761, 206-770, 206-773, 206-880, 209-949, 210-1034, 213-543, 218-699, 236-446, 255-763, 260-505, 262-948, 284-950, 307-1169, 324-939, 334-869, 399-1032, 436-924, 461-1079, 478-1108, 495-673, 496-730, 519-1171, 523-964, 523-1061, 527-705, 546-1137, 554-788, 566-1201, 567-1168, 570-1186, 570-1264, 571-1264, 591-884, 595-1185, 597-1055, 602-1211, 606-863, 612-1338, 622-1310, 624-1171, 630-1431, 631-1242, 652-1182, 652-1240, 663-1169, 668-1294, 690-1454, 699-1083, 705-1412, 711-1519, 713-1122, 713-1383, 720-953, 720-1025, 723-1280, 736-1225, 740-941, 744-1312, 764-1028, 768-1361, 772-1313, 772-1530, 780-1603, 780-1626, 783-1063, 785-1509, 800-1395, 809-1392, 813-1656, 824-1368, 826-1437, 831-1458, 834-1520, 839-1401, 844-1358, 853-1596, 855-1590, 857-1688, 858-1480, 869-1530, 873-1488, 885-1438, 886-1400, 888-1532, 896-1509, 914-1510, 915-1522, 917-1586, 922-1492,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	925-1616, 929-1462, 934-1538, 935-1409, 945-1638, 950-1679, 957-1542, 957-1626, 958-1666, 959-1438, 965-1628, 965-1659, 967-1512, 968-1509, 969-1409, 973-1467, 973-1733, 976-1669, 982-1875, 988-1542, 996-1573, 998-1667, 1000-1508, 1000-1601, 1010-1643, 1016-1431, 1016-1698, 1018-1741, 1019-1683, 1020-1590, 1022-1557, 1023-1483, 1026-1722, 1036-1626, 1039-1706, 1042-1728, 1043-1578, 1049-1715, 1050-1354, 1050-1431, 1058-1721, 1068-1778, 1075-1431, 1091-1660, 1091-1928, 1093-1924, 1106-1678, 1113-1672, 1114-1709, 1122-1848, 1124-1693, 1127-1769, 1137-1672, 1137-1733, 1137-1760, 1138-2032, 1143-1381, 1148-2023, 1154-1435, 1157-1662, 1162-1931, 1163-1617, 1172-2067, 1177-2017, 1181-1944, 1182-1808, 1192-1901, 1197-1494, 1197-1749, 1198-1673, 1199-1453, 1201-1947, 1207-1631, 1207-2104, 1208-2006, 1210-1459, 1213-1573, 1215-1461, 1215-1803, 1215-1860, 1220-1870, 1220-2085, 1228-1593, 1230-2085, 1232-1357, 1235-1826, 1235-1900, 1236-1522, 1242-1803, 1242-2064, 1246-2067, 1250-2085, 1251-2091, 1254-1542, 1260-1475, 1260-1617, 1260-2032,
	1265-2003, 1267-1438, 1271-1473, 1273-1593, 1273-1819, 1273-2111, 1276-1584, 1276-1617, 1287-1757, 1288-1936, 1290-1958, 1293-1546, 1294-1537, 1298-2087, 1301-1755, 1303-1537, 1304-1957, 1581-2111
55/ 3810039CB1/ 4017	1-647, 1-744, 104-857, 106-913, 111-834, 355-608, 483-1218, 619-1308, 630-1430, 816-1483, 816-1717, 962-1142, 962-1192, 962-1301, 962-1335, 962-1363, 962-1367, 962-1372, 962-1395, 962-1396, 962-1399, 963-1396, 991-1396, 1025-1396, 1120-1396, 1157-1396, 1164-1946, 1188-1396, 1267-1396, 1269-1396, 1284-1396, 1317-1626, 1317-1759, 1318-1648, 1823-2671, 1958-2679, 2138-2852, 2201-2660, 2201-2679, 2226-2858, 2266-2477, 2269-2683, 2552-3349, 2555-3349, 2608-3349, 2739-3272, 2900-3149, 3123-3592, 3123-3782, 3255-4017
56/ 8032337CB1/ 718	1-204, 1-335, 1-414, 1-466, 1-514, 1-532, 15-463, 51-414, 190-652, 217-694, 231-693, 238-690, 241-690, 248-691, 252-696, 257-691, 260-690, 261-690, 261-718, 314-690, 348-693, 365-691, 415-691, 505-690, 561-707
57/ 7506411CB1/ 4108	1-3998, 89-272, 89-351, 89-406, 89-413, 89-441, 89-489, 89-501, 89-517, 89-520, 90-492, 92-736, 94-606, 95-887, 100-517, 105-539, 107-527, 139-609, 211-899, 486-941, 493-927, 604-971, 611-965, 617-1281, 649-935, 678-1126, 679-1355, 694-969, 751-1552, 757-1485, 768-1008, 786-1480, 808-1042, 817-1413, 836-1486, 955-1591, 962-1591, 967-1591, 994-1486, 1015-1590, 1036-1670, 1095-1287, 1097-1417, 1097-1920, 1111-1512, 1119-1751, 1317-1944, 1356-1821, 1434-1710, 1479-1715, 1479-1739, 1481-2027, 1695-1937, 1759-2000, 1853-2101, 1884-2138, 1927-2475, 2027-2641, 2057-2350, 2061-2350, 2155-2406, 2251-2477, 2274-2722, 2280-2875, 2285-2621, 2287-2532, 2323-2544, 2334-2711, 2334-2714, 2363-2786, 2397-2654, 2462-2753, 2466-2781, 2556-2838, 2609-2855, 2631-2883, 2638-2872, 2640-2855, 2644-2882, 2654-2924, 2675-2905, 2716-3265, 2743-2871, 2743-3139, 2759-3047, 2766-3091, 2769-3091, 2774-3043, 2790-3059, 2792-2994, 2850-3104, 2856-3067, 2857-3153, 2872-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	3124, 2872-3137, 2903-3431, 2923-3184, 2923-3466, 2937-3131, 2943-3186, 2953-3183, 2964-3354, 2971-3231, 3008-3255, 3010-3300, 3042-3309, 3042-3332, 3063-3231, 3115-3379, 3135-3478, 3150-3435, 3170-3338, 3170-3432, 3172-3315, 3183-3447, 3195-3666, 3200-3674, 3210-3499, 3215-3471, 3215-3496, 3215-3765, 3233-3469, 3250-3497, 3251-3505, 3251-3741, 3252-3595, 3284-3552, 3294-3582, 3318-3527, 3318-3634, 3318-3881, 3328-3775, 3351-3624, 3351-3638, 3351-3753, 3363-3753, 3363-3762, 3388-3754, 3392-3753, 3409-3692, 3411-3753, 3412-3665, 3437-3573, 3442-3725, 3450-3753, 3453-3784, 3487-3753, 3487-3775, 3498-3753, 3518-3808, 3535-3758, 3554-3753, 3607-4025, 3607-4056, 3607-4108, 3648-3796, 3706-3975, 3731-3990, 3738-4093, 3740-3998, 3745-4003, 3753-4008, 3808-4108, 3876-3997
58/ 2658834CB1/ 6050	1-532, 1-543, 42-536, 47-752, 51-684, 125-644, 137-513, 173-771, 183-536, 208-643, 226-678, 226-818, 255-748, 277-783, 371-975, 397-1137, 453-740, 453-805, 453-865, 604-1245, 608-855, 660-1191, 837-1440, 1011-1235, 1067-1817, 1079-1282, 1092-1663, 1139-1683, 1158-1754, 1182-1694, 1182-1695, 1182-1708, 1182-1837, 1182-1950, 1182-1960, 1182-1982, 1182-1986, 1184-1694, 1185-1709, 1186-1924, 1230-1519, 1280-1880, 1458-2019, 1469-1655, 1571-2064, 1740-2002, 1740-2006, 1740-2146, 1768-2120, 1845-2336, 1873-2096, 1929-2360, 2097-2317, 2097-2527, 2097-2539, 2147-2693, 2217-2475, 2237-2489, 2261-2733, 2280-2543, 2376-2647, 2433-2927, 2501-2703, 2693-3014, 2813-3039, 2814-3092, 2900-5730, 4209-4319, 4210-4319, 4635-4667, 4635-4677, 4637-4667, 4767-4790, 4887-5825, 4904-5753, 4919-5769, 5003-5730, 5076-5186, 5121-5777, 5184-5988, 5194-5235, 5276-6000, 5285-5887, 5313-5989, 5332-5834, 5334-5946, 5352-5712, 5352-5928, 5352-5951, 5355-5778, 5371-
	6050, 5372-5957, 5409-5683, 5491-5787, 5508-5540, 5514-5778, 5535-5799, 5566-5834, 5593-5874, 5593-5881, 5599-5875, 5599-5929, 5629-5902, 5640-5663, 5752-5778
59/ 6818489CB1/ 3367	1-340, 1-395, 1-430, 1-450, 1-498, 1-508, 1-525, 1-530, 1-549, 1-574, 1-581, 1-607, 1-625, 1-3355, 2-630, 282-971, 341-993, 407-1099, 419-993, 517-950, 530-1183, 561-1273, 602-1367, 748-1296, 894-1595, 904-1607, 1033-1608, 1040-1527, 1042-1584, 1147-1938, 1163-1771, 1184-1720, 1213-1938, 1228-1870, 1229-1678, 1246-1929, 1250-1819, 1253-1808, 1256-1920, 1296-1877, 1307-1862, 1317-1907, 1349-1938, 1367-1917, 1370-1989, 1373-1907, 1434-2011, 1434-2136, 1438-1938, 1472-1888, 1521-2045, 1549-2039, 1565-2082, 1676-2309, 1682-1938, 1695-2325, 1695-2356, 1695-2373, 1695-2375, 1695-2378, 1695-2394, 1695-2395, 1695-2396, 1695-2439, 1695-2450, 1696-2426, 1697-2369, 1697-2371, 1700-2390, 1702-2498, 1727-2111, 1782-2183, 1867-2535, 1889-2420, 1909-2535, 1944-2535, 1982-2535, 1999-2535, 2010-2136, 2010-2300, 2010-2310, 2010-2336, 2010-2347, 2010-2423, 2010-2428, 2010-2489, 2010-2551, 2010-2556, 2010-2567, 2010-2571, 2010-2617, 2010-2621, 2010-2680, 2010-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	2718, 2010-2723, 2010-2813, 2010-2847, 2112-2535, 2120-2734, 2123-2666, 2124-2734, 2143-2668, 2171-2549, 2175-2561, 2203-2535, 2245-2387, 2258-2870, 2271-3050, 2330-2560, 2591-2954, 2614-3269, 2620-3267, 2620-3342, 2620-3367, 2708-3266, 2712-3332, 2713-3367, 2722-3271, 2726-3336, 2772-2955, 2782-2954, 2802-3366, 2818-2981, 2829-3055, 2831-3325, 2834-3337, 2868-3367, 2925-3089, 3078-3138
60/ 7509415CB1/ 2512	1-499, 1-597, 15-509, 505-1128, 505-1129, 508-1120, 508-1127, 508-1128, 509-785, 509-1128, 510-877, 510-881, 510-1007, 510-1014, 510-1029, 510-1038, 510-1101, 510-1123, 510-1128, 510-1129, 510-1173, 511-1128, 513-779, 513-783, 513-785, 513-960, 513-1110, 513-1121, 513-1131, 513-1133, 513-1181, 513-1187, 528-1128, 533-836, 535-1187, 541-785, 541-823, 541-858, 541-1128, 541-1181, 542-783, 550-1059, 552-763, 552-816, 552-849, 552-913, 552-1026, 552-1029, 552-1052, 552-1053, 552-1055, 552-1092, 552-1143, 552-1185, 552-1187, 559-1126, 562-1128, 903-945, 903-953, 1015-1350, 1029-1600, 1029-1715, 1043-1545, 1062-1184, 1062-1187, 1130-1364, 1452-1950, 1683-2443, 1683-2494, 2050-2512
61/ 7506916CB1/ 2147	1-558, 298-1194, 326-851, 326-1769, 470-1285, 470-1290, 471-968, 471-1012, 471-1014, 471-1019, 471-1061, 471-1109, 471-1126, 471-1209, 471-1285, 478-1212, 488-1189, 492-740, 492-909, 492-944, 492-963, 492-987, 492-1032, 492-1106, 532-914, 602-843, 644-1020, 659-1173, 699-952, 699-1164, 721-1557, 723-1289, 735-1056, 763-1270, 779-1178, 781-1094, 781-1260, 781-1274, 782-1269, 854-1176, 856-1290, 883-1018, 905-1268, 966-1290, 971-1183, 1183-1286, 1325-1732, 1339-1688, 1339-1691, 1352-1688, 1387-1720, 1422-2147, 1455-1698, 1455-1779, 1457-1666, 1460-1561, 1461-1703, 1484-1752, 1502-1712, 1502-1769, 1502-1773, 1510-1639, 1557-1762, 1778-2079
62/ 7507104CB1/ 1421	1-578, 1-1414, 181-459, 258-487, 284-1154, 401-1320, 599-825, 602-875, 639-1214, 670-966, 690-1258, 734-987, 734-1133, 748-991, 774-1415, 782-1374, 976-1420, 1011-1421, 1018-1414, 1020-1266, 1069-1418, 1075-1418, 1103-1418, 1129-1420, 1141-1415, 1164-1419, 1181-1413, 1218-1421, 1230-1418
63/ 7507105CB1/ 1481	1-1473, 181-459, 258-487, 284-1046, 524-758, 658-884, 661-934, 698-1273, 729-1025, 749-1317, 793-1046, 793-1192, 807-1050, 833-1474, 841-1433, 1035-1479, 1070-1481, 1077-1473, 1079-1325, 1128-1477, 1134-1477, 1162-1477, 1188-1479, 1200-1474, 1223-1478, 1240-1472, 1277-1481, 1289-1477
64/ 7507107CB1/ 1472	1-617, 1-1463, 181-459, 239-745, 258-487, 270-795, 284-1211, 400-640, 400-876, 455-1369, 560-812, 1025-1469, 1060-1471, 1067-1463, 1069-1315, 1118-1467, 1124-1467, 1152-1467, 1178-1469, 1190-1464, 1213-1468, 1230-1462, 1267-1472, 1279-1467
65/ 7507109CB1/ 1480	1-617, 1-1471, 181-459, 239-745, 258-487, 270-795, 284-1097, 400-640, 400-876, 560-812, 673-1377, 712-938, 715-988, 783-1079, 847-1100, 861-1104, 1161-1475, 1186-1477, 1198-1472, 1221-1476, 1238-1470, 1275-1480, 1287-1475
66/ 1833937CB1/ 515	1-113, 1-226, 1-264, 1-460, 1-497, 1-500, 1-502, 1-515, 3-490, 3-515, 82-425, 346-515, 385-515

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
67/ 7502036CB1/ 1961	1-177, 1-179, 1-235, 1-252, 1-260, 1-273, 1-512, 1-513, 4-144, 6-233, 7-536, 25-273, 34-273, 87-177, 115-177, 178-790, 192-725, 193-771, 251-755, 251-756, 261-856, 272-493, 272-857, 274-857, 275-756, 275-857, 281-756, 284-855, 289-852, 289-857, 290-857, 302-756, 302-757, 317-601, 317-818, 370-956, 377-838, 382-820, 385-930, 395-857, 412-978, 424-748, 432-695, 457-702, 458-729, 475-822, 475-925, 475-981, 491-719, 500-758, 528-780, 538-1026, 555-1059, 580-857, 612-1138, 625-1135, 631-772, 645-1282, 649-857, 696-857, 702-966, 748-1131, 1041-1216, 1041-1590, 1055-1657, 1092-1327, 1118-1297, 1186-1317, 1186-1427, 1193-1483, 1220-1445, 1305-1669, 1310-1548, 1337-1425, 1345-1644, 1371-1611, 1394-1648, 1402-1617, 1405-1601, 1470-1716, 1471-1744, 1476-1737, 1480-1761, 1497-1725, 1507-1750, 1507-1770, 1517-1743, 1525-1732, 1528-1813, 1529-1904, 1547-1779, 1657-1961
68/ 7503248CB1/ 1154	1-190, 1-193, 1-194, 1-195, 1-315, 1-396, 1-596, 1-1102, 8-335, 8-612, 21-195, 22-505, 22-618, 23-530, 23-629, 25-352, 25-646, 35-195, 36-195, 40-403, 40-469, 40-509, 40-515, 40-654, 52-195, 58-654, 95-623, 122-258, 122-784, 122-786, 122-787, 148-785, 169-307, 169-530, 169-650, 187-652, 201-637, 283-498, 286-654, 287-624, 292-635, 305-585, 956-1077, 964-1078, 979-1154, 982-1081, 992-1141
69/ 7503968CB1/ 1938	1-262, 1-1929, 1-1936, 10-388, 79-325, 109-285, 109-290, 109-323, 109-537, 109-582, 110-285, 112-544, 132-285, 1063-1599, 1246-1913, 1285-1914, 1316-1758, 1316-1824, 1316-1849, 1316-1870, 1316-1934, 1316-1935, 1316-1936, 1316-1938, 1321-1936, 1323-1936, 1327-1936, 1334-1936, 1383-1891, 1454-1936, 1492-1824, 1492-1848, 1492-1915, 1492-1927, 1492-1928, 1492-1929, 1492-1930, 1492-1935, 1492-1936, 1492-1938, 1496-1929, 1498-1936, 1592-1914
70/ 7505931CB1/ 931	1-141, 1-229, 1-230, 1-261, 1-282, 1-293, 1-304, 1-429, 1-931, 3-275, 18-437, 48-334, 48-581, 48-583, 50-339, 96-356, 428-581, 478-911, 588-838, 625-871, 682-921
71/ 7506912CB1/ 1730	1-749, 1-918, 1-923, 1-947, 1-948, 1-1730, 144-893, 240-541, 240-620, 242-614, 242-806, 311-1308, 357-641, 357-858, 357-1035, 357-1085, 357-1169, 368-1098, 388-1308, 396-1308, 399-1308, 409-1308, 420-1308, 424-1308, 466-1308, 474-1308, 475-1307, 481-1307, 509-1108, 523-716, 524-689, 524-1308, 556-1308, 560-1308, 560-1443, 595-1308, 597-1228, 601-1307, 602-1308, 627-879, 627-1123, 647-1308, 685-1308, 825-1101, 831-1308, 869-1072, 876-1346, 889-1488, 891-1137, 891-1316, 893-1262, 905-1192, 907-1377, 923-1178, 991-1630, 996-1650, 1019-1286, 1055-1705, 1072-1597, 1082-1692, 1085-1648, 1147-1730, 1175-1430, 1204-1730, 1268-1491, 1333-1730, 1334-1570, 1334-1605, 1576-1730
72/ 7506913CB1/ 1495	1-833, 1-1495, 141-1072, 182-483, 182-562, 184-556, 281-1028, 299-583, 299-630, 330-986, 408-1073, 465-635, 466-631, 628-837, 641-1111, 654-1253, 656-902, 656-1081, 658-1027, 670-957, 672-1142, 688-943, 756-1395, 761-1415, 784-1051, 820-1470, 837-1362, 847-1457, 850-1413, 912-1495, 937-1195, 969-1495, 1033-1256, 1098-1495, 1099-1335, 1099-1370, 1341-1495
73/ 7507029CB1/ 1153	1-561, 1-1153, 125-585, 245-421, 245-585, 275-578, 358-694, 358-920, 403-668, 689-1141, 709-1142, 760-1141, 852-1141, 914-1153, 924-1153, 966-1151

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
74/ 7507063CB1/ 1472	1-127, 7-126, 7-127, 7-1472, 9-126, 14-380, 14-796, 23-127, 24-126, 24-127, 277-1094, 379-1084, 379-1094, 379-1095, 379-1099, 382-1094, 386-1094, 403-1094, 411-1095, 412-1094, 413-1094, 426-570, 426-1094, 428-1094, 429-1094, 433-1094, 470-1094, 474-703, 485-1099, 490-1117, 495-1094, 500-1091, 500-1097, 503-1091, 504-1094, 522-1094, 527-1091, 529-1083, 582-845, 583-882, 584-864, 585-1168, 591-832, 591-1186, 599-828, 599-1299, 601-1087, 608-1087, 622-899, 651-1322, 659-1109, 662-1215, 672-1257, 692-1292, 697-979, 733-882, 733-983, 740-1322, 741-1209, 759-1115, 810-1303, 842-1005, 866-1094, 912-1426, 931-1426, 996-1255, 1164-1427
75/ 7504755CB1/ 2997	1-2997, 1140-1346, 1241-2024, 1241-2107, 1354-1841, 1354-1868, 1354-1887, 1354-1889, 1354-1912, 1354-2000, 1354-2021, 1354-2071, 1354-2121, 1354-2125, 1356-2128, 1359-2111, 1427-2217, 1442-1670, 1442-1713, 1442-2102, 1443-2102, 1446-2122, 1460-2102, 1463-2102, 1474-2011, 1479-2180, 1518-2102, 1533-2102, 1537-2102, 1559-2160, 1573-2084, 1589-2244, 1599-2226, 1615-2177, 1629-2102, 1656-2102, 1662-2102, 1711-2102, 1713-2427, 1717-2390, 1732-1894, 1763-2102, 1794-2102, 1817-2493, 1821-2396, 1823-2449, 1835-2438, 1847-2106, 1862-2462, 1869-2102, 1875-2102, 1891-2762, 1909-2442, 1970-2598, 1983-2560, 2007-2516, 2020-2933, 2027-2933, 2045-2267, 2048-2933, 2076-2549, 2099-2946, 2104-2375, 2116-2287, 2123-2933, 2149-2384, 2190-2845, 2222-2933, 2247-2851, 2259-2851, 2259-2892, 2264-2887, 2278-2864, 2284-2595, 2289-2917, 2304-2929, 2331-2933, 2360-2929, 2367-2946, 2376-2946, 2403-2620, 2403-2918, 2410-2929, 2425-2930, 2428-2622, 2428-
	2724, 2428-2850, 2428-2925, 2428-2946, 2440-2946, 2467-2933, 2482-2714, 2482-2915, 2482-2946, 2488-2946, 2493-2946, 2495-2774, 2500-2949, 2535-2946, 2550-2946, 2584-2928, 2616-2928, 2634-2946, 2688-2901, 2809-2946
76/ 7509265CB1/ 2046	1-2019, 70-322, 70-362, 146-315, 260-1055, 420-974, 420-1188, 420-1259, 433-1259, 475-637, 489-1259, 495-1258, 498-1259, 508-804, 534-1259, 582-727, 582-733, 607-1259, 653-829, 843-1333, 846-993, 846-1067, 846-1105, 846-1259, 846-1365, 846-1422, 846-1428, 857-1362, 862-1336, 862-1411, 867-1396, 868-1559, 872-1431, 897-1508, 899-1491, 908-1508, 914-1478, 917-1390, 919-1127, 927-1458, 950-1504, 972-1431, 977-1613, 989-1159, 989-1558, 999-2019, 1001-1517, 1002-1259, 1004-1479, 1008-1403, 1013-1553, 1014-1423, 1014-1576, 1023-1602, 1024-1503, 1024-1619, 1026-1227, 1032-1453, 1035-1614, 1039-1509, 1040-1511, 1042-1271, 1042-1482, 1045-1274, 1047-1521, 1052-1281, 1052-1333, 1056-1548, 1063-1639, 1064-1546, 1067-1775, 1069-1622, 1074-1602, 1080-1238, 1082-1338, 1090-2019, 1092-1570, 1095-1689, 1096-1748, 1108-2019, 1116-1658, 1121-1389, 1128-1602, 1133-2019, 1134-2019, 1136-1442, 1137-1848, 1140-1411, 1144-1263, 1144-1644, 1145-2019, 1146-1672, 1156-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
	1679, 1158-1658, 1165-1754, 1171-1675, 1174-1404, 1181-1789, 1184-1774, 1186-1630, 1187-1663, 1189-2019, 1190-1658, 1192-2019, 1203-1801, 1207-1721, 1210-1493, 1215-1802, 1217-2019, 1233-1862, 1239-1438, 1241-1674, 1241-1733, 1252-1774, 1259-1760, 1259-1796, 1272-1760, 1276-1854, 1279-1664, 1293-1504, 1319-1633, 1331-1625, 1342-2019, 1349-1984, 1351-1719, 1351-1907, 1351-1991, 1354-1841, 1357-1791, 1365-1848, 1383-1651, 1388-1686, 1407-1726, 1419-2019, 1434-1664, 1434-1675, 1434-2025, 1451-1981, 1468-1680, 1502-2046, 1504-2039, 1505-2009, 1520-1760, 1524-1811, 1543-2013, 1552-2014, 1556-2012, 1560-1833, 1561-1788, 1574-1818, 1614-1753, 1615-1832, 1623-2026, 1624-2012, 1689-1912, 1754-1973, 1765-2017, 1779-2013, 1780-2034, 1789-1941, 1803-2024, 1804-2028
77/ 7509371CB1/ 1668	1-217, 1-882, 1-1668, 2-217, 19-144, 19-191, 19-291, 19-432, 21-676, 21-677, 21-679, 21-689, 21-727, 21-728, 21-739, 21-741, 21-779, 21-807, 21-808, 21-839, 21-840, 21-870, 21-890, 21-894, 21-901, 37-663, 37-678, 37-716, 37-726, 37-752, 37-808, 37-816, 37-897, 37-953, 37-990, 52-217, 84-217, 87-217, 112-217, 113-214, 153-217, 156-217, 157-218, 204-319, 232-611, 692-815, 692-816, 762-1490, 1345-1662
78/ 7509389CB1/ 1735	1-217, 1-1630, 2-217, 19-144, 19-191, 19-291, 19-432, 21-758, 37-938, 52-217, 84-217, 87-217, 112-214, 112-217, 153-217, 156-217, 157-217, 157-218, 204-319, 232-598, 609-1518, 615-1250, 673-1518, 678-1518, 722-1518, 746-1518, 764-1295, 776-1518, 778-1518, 797-1518, 813-1518, 815-1518, 864-1518, 873-1518, 878-1518, 904-1518, 919-1163, 919-1479, 919-1535, 920-1518, 949-1518, 961-1518, 1021-1518, 1075-1336, 1076-1353, 1373-1735
79/ 7507005CB1/ 1413	1-270, 1-1413, 4-117, 4-218, 4-250, 56-284, 97-340, 112-219, 112-523, 112-680, 112-711, 112-725, 112-788, 112-803, 112-804, 112-807, 112-809, 112-810, 112-811, 112-885, 112-888, 112-889, 112-894, 112-914, 112-928, 112-965, 112-1012, 113-805, 113-950, 123-402, 134-577, 268-772, 286-768, 293-768, 295-559, 295-773, 310-747, 327-1020, 367-780, 370-436, 376-440, 379-1328, 382-758, 395-1328, 406-1327, 406-1328, 443-1328, 448-1328, 478-1328, 495-1328, 509-1328, 510-1328, 564-1328, 566-1328, 572-1327, 572-1328, 573-1328, 585-1328, 586-1328, 594-1168, 594-1328, 604-1328, 629-1170, 633-1328, 634-1328, 656-1328, 660-1328, 674-1328, 675-1328, 782-1175, 818-938, 818-1200, 818-1262, 880-1139, 940-1174, 967-1244, 994-1412, 1079-1413
80/ 7509142CB1/ 935	1-208, 1-258, 1-374, 1-401, 1-456, 1-465, 1-488, 1-498, 1-530, 1-541, 1-580, 1-591, 8-580, 80-601, 119-749, 121-675, 145-749, 154-591, 156-935, 162-745, 206-728, 224-672, 250-746, 274-422, 426-631
81/ 7509157CB1/ 2038	1-221, 169-923, 327-475, 757-1351, 764-1085, 812-1653, 825-1654, 831-1656, 832-1653, 873-1656, 890-1653, 928-1654, 928-1656, 938-1653, 962-1562, 965-1653, 969-1501, 994-1625, 994-1654, 1003-1575, 1065-1546, 1109-1716, 1114-1590, 1151-1736, 1156-1709, 1210-1892, 1233-1489, 1235-1892, 1237-1768, 1237-1867, 1238-1675, 1257-1842, 1279-1720, 1312-1892, 1321-1892, 1332-1500, 1352-1506, 1362-1892, 1373-1529, 1373-1890, 1431-1892, 1443-1892, 1471-1892, 1493-1892, 1520-1892, 1539-1775, 1539-2038, 1550-1892, 1577-1892, 1660-1973, 1733-1959

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
82/ 7509246CB1/ 1471	1-127, 7-126, 7-127, 7-1471, 14-828, 23-127, 24-126, 24-127, 387-1094, 426-570, 474-703, 490-1117, 529-1083, 582-845, 583-882, 584-864, 585-1168, 591-832, 591-1186, 599-828, 599-1299, 601-1087, 608-1087, 622-899, 651-1322, 659-1109, 662-1215, 672-1257, 692-1292, 697-979, 733-882, 733-983, 740-1322, 741-1209, 759-1115, 810-1303, 842-1005, 912-1425, 931-1313, 996-1255, 1164-1426
83/ 7509380CB1/ 1601	1-221, 2-604, 5-554, 14-271, 14-469, 14-501, 14-511, 14-543, 14-593, 21-593, 93-614, 132-762, 134-688, 158-762, 167-604, 169-964, 169-972, 169-975, 169-1003, 169-1061, 169-1063, 169-1067, 175-758, 195-820, 219-741, 237-685, 263-759, 287-435, 319-860, 349-1096, 359-852, 380-931, 393-920, 420-931, 430-931, 439-644, 473-1008, 788-1601, 819-1601
84/ 7509382CB1/ 1404	1-221, 14-271, 132-762, 176-237, 236-850, 237-685, 263-759, 287-435, 439-644, 462-1165, 786-1353, 788-1398, 790-1231, 823-1404, 832-1404, 843-1011, 863-1017, 873-1404, 884-1040, 884-1401, 942-1404, 954-1404, 982-1404, 1004-1404, 1031-1404, 1050-1286, 1088-1404

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
43	7503679CB1	BRAFNOT02
44	7503681CB1	BRAFNOT02
45	7505819CB1	THYMNOR02
46	7505083CB1	UCMCL5T01
47	7505866CB1	HNT2NOT01
48	7503214CB1	BRADDIR01
49	7495312CB1	MUSCDIT06
50	7506732CB1	COLNNOT19
51	7506736CB1	OVARNOT12
52	7507121CB1	PLACFER01
54	1967990CB1	BRSTNOT04
55	3810039CB1	BRADDIR01
56	8032337CB1	TESTNOF01
57	7506411CB1	MIXDTXE01
58	2658834CB1	SEMVTDE01
59	6818489CB1	BRAUNOR01
60	7509415CB1	BMARTXE01
61	7506916CB1	BRAINOT11
62	7507104CB1	NGANNOT01
63	7507105CB1	NGANNOT01
64	7507107CB1	COLNTUT03
65	7507109CB1	COLNTUT03
66	1833937CB1	BRAINON01
67	7502036CB1	BRACNOK02
68	7503248CB1	THYMNOR02
69	7503968CB1	UTRSTMR02
70	7505931CB1	EOSITXT01
71	7506912CB1	STOMFET02
72	7506913CB1	BRADDIR01
73	7507029CB1	DRGCNOT02
74	7507063CB1	UTRSTMR01
75	7504755CB1	SYNORAT03
76	7509265CB1	BRAIFEN08
77	7509371CB1	KIDNNOT19
78	7509389CB1	KIDNNOT19
79	7507005CB1	BRAVUNT02
80	7509142CB1	BRADDIR01
81	7509157CB1	BRADDIR01
82	7509246CB1	UTRSTMR01
83	7509380CB1	BRADDIR01
84	7509382CB1	BRADDIR01

Table 6

Library	Vector	Library Description
BMARTXE01	pINCY	This 5' biased random primed library was constructed using RNA isolated from treated SH-SY5Y cells derived from a metastatic bone marrow neuroblastoma, removed from a 4-year-old Caucasian female (Schering AG). The medium was MEM/HAM'S F12 with 10% fetal calf serum. After reaching about 80% confluency cells were treated with 6-Hydroxydopamine (6-OHDA) at 100 microM for 8 hours.
BRACNOK02	PSPORT1	This amplified and normalized library was constructed using RNA isolated from posterior cingulate tissue removed from an 85-year-old Caucasian female who died from myocardial infarction and retroperitoneal hemorrhage. Pathology indicated atherosclerosis, moderate to severe, involving the circle of Willis, middle cerebral, basilar and vertebral arteries; infarction, remote, left dentate nucleus; and amyloid plaque deposition consistent with age. There was mild to moderate leptomeningeal fibrosis, especially over the convexity of the frontal lobe. There was mild generalized atrophy involving all lobes. The white matter was mildly thinned. Cortical thickness in the temporal lobes, both maximal and minimal, was slightly reduced. The substantia nigra pars compacta appeared mildly depigmented. Patient history included COPD, hypertension, and recurrent deep venous thrombosis. 6.4 million independent clones from this amplified library were normalized in one round using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791.
BRADDIR01	pINCY	Library was constructed using RNA isolated from diseased choroid plexus tissue of the lateral ventricle, removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.
BRAFNOT02	pINCY	Library was constructed using RNA isolated from superior frontal cortex tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. In addition, scattered throughout the cerebral cortex, there were multiple small microscopic areas of cavitation with surrounding gliosis. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly, and an enlarged spleen and liver.
BRAIFEN08	pINCY	This normalized fetal brain tissue library was constructed from 400 thousand independent clones from a fetal brain tissue library. Starting RNA was made from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6

Library	Vector	Library Description
BRAINON01	PSPORT1	Library was constructed and normalized from 4.88 million independent clones from a brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:9228, except that a significantly longer (48-hour) reannealing hybridization was used.
BRAINOT11	pINCY	Library was constructed using RNA isolated from brain tissue removed from the right temporal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), consistent with chronic seizure disorder. Family history included a cervical neoplasm.
BRAUNOR01	pINCY	This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex.
		There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.
BRAVUNT02	PSPORT1	Library was constructed using pooled RNA isolated from separate populations of unstimulated astrocytes.
BRSTNOT04	PSPORT1	Library was constructed using RNA isolated from breast tissue removed from a 62-year-old East Indian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated an invasive grade 3 ductal carcinoma. Patient history included benign hypertension, hyperlipidemia, and hematuria. Family history included cerebrovascular and cardiovascular disease, hyperlipidemia, and liver cancer.
COLNNOT19	pINCY	Library was constructed using RNA isolated from the cecal tissue of an 18-year-old Caucasian female. The cecal tissue, along with the appendix and ileum tissue, were removed during bowel anastomosis. Pathology indicated Crohn's disease of the ileum involving 15 cm of the small bowel.

Table 6

Library	Vector	Library Description
COLNTUT03	pINCY	Library was constructed using RNA isolated from colon tumor tissue obtained from the sigmoid colon of a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy. Pathology indicated invasive grade 2 adenocarcinoma. One lymph node contained metastasis with extranodal extension. Patient history included hyperlipidemia, cataract disorder, and dermatitis. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, breast cancer, and prostate cancer.
DRGCNOT02	pINCY	Library was constructed using RNA isolated from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema, acute bronchopneumonia, bilateral pleural effusions, pericardial effusion, and malignant lymphoma (natural killer cell type). The patient presented with pyrexia of unknown origin, malaise, fatigue, and gastrointestinal bleeding. Patient history included probable cytomegalovirus infection, liver congestion, and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia of the left lung, natural killer cell lymphoma of the pharynx, Bell's palsy, and tobacco and alcohol abuse. Previous surgeries included colonoscopy, closed colon biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy. Patient medications included Diflucan (fluconazole), Deltasone (prednisone), hydrocodone, Lortab, Alprazolam, Reazodone, ProMace-Cytabom, Etoposide, Cisplatin, Cytarabine, and dexamethasome. The patient received radiation therapy and multiple blood transfusions.
EOSITXT01	pINCY	Library was constructed using RNA isolated from eosinophils stimulated with IL-5.
HNT2NOT01	PBLUESCRIPT	Library was constructed at Stratagene (STR937230), using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor).
KIDNNOT19	pINCY	Library was constructed using RNA isolated from kidney tissue removed a 65-year-old Caucasian male during an exploratory laparotomy and nephroureterectomy. Pathology for the associated tumor tissue indicated a grade 1 renal cell carcinoma within the upper pole of the left kidney. Patient history included malignant melanoma of the abdominal skin, benign neoplasm of colon, cerebrovascular disease, and umbilical hernia. Family history included myocardial infarction, atherosclerotic coronary artery disease, cerebrovascular disease, prostate cancer, myocardial infarction, and atherosclerotic coronary artery disease.
MIXDTXE01	PBK-CMV	This 5' biased random primed library was constructed using pooled cDNA from nine donors. cDNA was generated using mRNA isolated from Jurkat cell line derived from the T cells of a male (donor A), THP-1 cell line derived from the peripheral blood of a 1-year-old male (donor B), Daudi cell line derived from B-lymphoblasts from a 16-year-old black male (donor C), RPMI-1666 cell line derived from lymphoma tissue from a 29-year-old Caucasian male (donor D), spleen from a 1-year-old Caucasian male (donor E), thymus removed from a 21-year-old Caucasian male (donor F) during a thymectomy, lymph node from a 42-year-old Caucasian female (donor G), thymus tumor from a 56-year-old Caucasian female (donor H) during a total thymectomy and PBMC's from a pool of donors (donor I). The patients presented with anemia and persistent hyperplastic thymus (H). Patient history included acute T-cell leukemia (A); acute monocytic leukemia (B); Burkitt's lymphoma (C); Hodgkin's disease (D); Bronchitis (E); hydrocele, regional enteritis or the small intestine, atopic dermatitis and benign

Table 6

Library	Vector	Library Description
		and cardiac dysrhythmia and left bundle branch block (H). Previous surgeries included an appendectomy and parathyroid surgery (F); unspecified heart surgery (G); and a normal delivery (H). Family history included benign hypertension in the grandparent(s) and coronary artery disease in the father of donor F.
MUSCDIT06	pINCY	Library was constructed using RNA isolated from skeletal muscle tissue removed from an 11-month-old Caucasian female who died from cardiopulmonary arrest. Patient history included Krabbe's disease.
NGANNOT01	PSPORT1	Library was constructed using RNA isolated from tumorous neuroganglion tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma. Family history included asthma.
OVARNOT12	pINCY	Library was constructed using RNA isolated from ovarian tissue removed from a 36-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, soft tissue excision, and an incidental appendectomy. Pathology for the associated tumor tissue indicated one intramural and one subserosal leiomyata of the myometrium. The cervix showed mild chronic cervicitis. A left inner thigh excision showed a lipoma. Family history included hyperlipidemia, acute myocardial infarction, atherosclerotic coronary artery disease, type II diabetes, and chronic liver disease.
PLACFER01	pINCY	The library was constructed using RNA isolated from placental tissue removed from a Caucasian fetus, who died after 16 weeks' gestation from fetal demise and hydrocephalus. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time). Serology was positive for anti-CMV. Family history included multiple pregnancies and live births, and an abortion.
SEMVTDE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from seminal vesicle tissue removed from a 63-year-old Caucasian male during closed prostatic biopsy, radical prostatectomy, and regional lymph node excision. Pathology for the associated tumor tissue indicated Gleason grade 2+3 adenocarcinoma in the right side of the prostate. Adenofibromatous hyperplasia was present. The patient presented with prostate cancer, elevated prostate specific antigen and prostatic hyperplasia. Patient history included kidney calculus, extrinsic asthma, benign bowel neoplasm, backache, tremor, and tobacco abuse in remission. Previous surgeries included adenotonsillectomy. Patient medications included Ventolin and Vanceril. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and acute myocardial infarction in the father; and stomach cancer and extrinsic asthma in the grandparent(s).
STOMFET02	pINCY	Library was constructed using RNA isolated from stomach tissue removed from a Hispanic male fetus, who died at 18 weeks' gestation.
SYNORAT03	PSPORT1	Library was constructed using RNA isolated from the wrist synovial membrane tissue of a 56-year-old female with rheumatoid arthritis.

Table 6

Library	Vector	Library Description
TESTNOF01	PSPORT1	This 5' cap isolated full-length library was constructed using RNA isolated from testis tissue removed from a 26-year-old Caucasian male who died from head trauma due to a motor vehicle accident. Serologies were negative. Patient history included a hernia at birth, tobacco use (1 1/2 ppd), marijuana use, and daily alcohol use (beer and hard liquor).
THYMNOR02	pINCY	The library was constructed using RNA isolated from thymus tissue removed from a 2-year-old Caucasian female during a thymectomy and patch closure of left atrioventricular fistula. Pathology indicated there was no gross abnormality of the thymus. The patient presented with congenital heart abnormalities. Patient history included double inlet left ventricle and a rudimentary right ventricle, pulmonary hypertension, cyanosis, subaortic stenosis, seizures, and a fracture of the skull base. Family history included reflux neuropathy.
UCMCL5T01	PBLUESCRIPT	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of 12 individuals. The cells were cultured for 12 days with IL-5 before RNA was obtained from the pooled lysates.
UTRSTMR01	pINCY	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy. The endometrium was secretory and contained fragments of endometrial polyps. Pathology for associated tumor tissue indicated uterine leiomyoma. Patient history included ventral hernia and a benign ovarian neoplasm.
UTRSTMR02	PCDNA2.1	This random primed library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from endometrial tissue removed from a 32-year-old female (donor A) and using mRNA isolated from myometrium removed from a 45-year-old female (donor B) during vaginal hysterectomy and bilateral salpingo-oophorectomy. In donor A, pathology indicated the endometrium was secretory phase. The cervix showed severe dysplasia (CIN III) focally involving the squamocolumnar junction at the 1, 6 and 7 o'clock positions. Mild koilocytotic dysplasia was also identified within the cervix. In donor B, pathology for the matched tumor tissue indicated multiple (23) subserosal, intramural, and submucosal leiomyomata. Patient history included stress incontinence, extrinsic asthma without status asthmaticus and normal delivery in donor B. Family history included cerebrovascular disease, depression, and atherosclerotic coronary artery disease in donor B.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value = 1.0E-8 or less; Full Length sequences: Probability value = 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value = 1.06E-6; Assembled ESTs: fasta Identity = 95% or greater and Match length = 200 bases or greater; fastx E value = 1.0E-8 or less; Full Length sequences: fastx score = 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-	Probability value = 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM, INCY, SMART and TIGRFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-	PFAM, INCY, SMART or TIGRFAM hits: Probability value = 1.0E-3 or less; Signal peptide hits: Score = 0 or greater
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score ≥ GCG specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.

Table 7

Program	Description	Reference	Parameter Threshold
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score = 3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
53	90086258	1629982H1	SNP00104067	190	1160	T	T	C	noncoding	n/d	n/d	0.98	n/d
53	90086258	1811361H1	SNP00136838	45	411	A	A	G	K126	n/a	n/a	n/a	n/a
53	90086258	1955494H1	SNP00136836	179	64	C	C	T	P10	n/a	n/a	n/a	n/a
53	90086258	3092405H1	SNP00033096	184	1248	C	C	T	noncoding	n/a	n/a	n/a	n/a
53	90086258	4161676H1	SNP00136837	223	76	G	G	A	G14	n/a	n/a	n/a	n/a
53	90086258	7050550H1	SNP00104068	396	1212	G	G	A	noncoding	n/d	n/a	n/a	n/a
55	3810039	8060531J1	SNP00147601	267	3701	A	G	A	noncoding	n/a	n/a	n/a	n/a
57	7506411	1805287H1	SNP00012602	83	3053	T	T	C	noncoding	0.82	n/a	n/a	n/a
57	7506411	1914405H1	SNP00045835	138	3490	C	C	G	noncoding	n/a	n/a	n/a	n/a
57	7506411	2203554H1	SNP00045834	155	3405	G	G	A	noncoding	n/a	n/a	n/a	n/a
57	7506411	4959088H1	SNP00045833	37	856	G	G	A	R245	n/a	n/a	n/a	n/a
58	2658834	2675350H1	SNP00058983	79	3987	C	C	A	H1218	n/a	n/a	n/a	n/a
58	2658834	2718577H1	SNP00062739	70	4810	C	C	T	L1493	n/a	n/a	n/a	n/a
58	2658834	2718577H1	SNP00069753	93	4833	C	C	T	G1500	n/d	n/a	n/a	n/a
58	2658834	2718577H1	SNP00141271	47	4790	A	A	G	Y1486	n/a	n/a	n/a	n/a
58	2658834	2809047H1	SNP00141272	37	5791	C	C	T	noncoding	n/a	n/a	n/a	n/a
58	2658834	3333543H1	SNP00069753	26	5709	C	C	T	D1792	n/d	n/a	n/a	n/a
58	2658834	3333543H1	SNP00108814	31	5714	A	A	G	Y1794	0.76	n/a	n/a	n/a
58	2658834	3541113H1	SNP00062739	7	5686	C	C	T	L1785	n/a	n/a	n/a	n/a
58	2658834	3697148H1	SNP00029988	145	5163	C	C	T	D1610	n/a	n/a	n/a	n/a
58	2658834	3747046H1	SNP00144771	197	3591	C	T	C	T1086	n/a	n/a	n/a	n/a
58	2658834	6283214H1	SNP00141271	62	5663	A	A	G	Y1777	n/a	n/a	n/a	n/a
58	2658834	6286171H2	SNP00062739	88	4811	T	C	T	L1493	n/a	n/a	n/a	n/a
58	2658834	6287256H2	SNP00108814	116	4837	A	A	G	T1502	0.76	n/a	n/a	n/a
58	2658834	6468436H1	SNP00029987	100	3411	G	G	A	V1026	n/a	n/a	n/a	n/a
58	2658834	6526679H1	SNP00058983	118	4854	C	C	A	R1507	n/a	n/a	n/a	n/a
58	2658834	6763311H1	SNP00144772	238	3639	C	T	C	P1102	n/a	n/a	n/a	n/a
58	2658834	6806251H1	SNP00058983	285	4853	C	C	A	P1507	n/a	n/a	n/a	n/a
58	2658834	6871656H1	SNP00006632	138	4710	C	T	C	H1459	0.84	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
58	2658834	6871656H1	SNP00108814	272	4843	A	A	G	M1504	0.76	n/a	n/a	n/a
58	2658834	7044804H1	SNP00006632	5	5583	T	T	C	Y1750	0.84	n/a	n/a	n/a
58	2658834	7638887H1	SNP00062739	567	4813	C	C	T	L1494	n/a	n/a	n/a	n/a
60	7509415	7278940H1	SNP00139037	297	1656	C	C	T	V389	n/a	n/a	n/a	n/a
66	1833937	131900H1	SNP00043506	75	84	T	G	T	I21	n/d	n/a	n/a	n/a
66	1833937	1701727H1	SNP00051244	166	163	T	T	G	I47	n/a	n/a	n/a	n/a
66	1833937	2563732H1	SNP00146052	245	329	T	T	G	noncoding	n/a	n/a	n/a	n/a
66	1833937	2648731H1	SNP00043506	57	67	G	G	T	L15	n/d	n/a	n/a	n/a
66	1833937	2648731H1	SNP00072770	154	164	G	G	A	G48	n/a	n/a	n/a	n/a
66	1833937	2838439H2	SNP00032201	47	458	C	C	T	noncoding	n/a	n/a	n/a	n/a
66	1833937	2838439H2	SNP00142786	202	303	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	1833937	3049613H1	SNP00072770	122	147	G	G	A	C42	n/a	n/a	n/a	n/a
66	1833937	3073391H1	SNP00051244	41	164	T	T	G	C48	n/a	n/a	n/a	n/a
66	1833937	3130302H1	SNP00043506	60	85	G	G	T	R21	n/d	n/a	n/a	n/a
66	1833937	3171639H1	SNP00043506	100	73	G	G	T	S17	n/d	n/a	n/a	n/a
66	1833937	3205405H1	SNP00051244	144	162	T	T	G	I47	n/a	n/a	n/a	n/a
66	1833937	3540827H1	SNP00043506	77	83	G	G	T	G21	n/d	n/a	n/a	n/a
66	1833937	3558864H1	SNP00043506	69	81	G	G	T	R20	n/d	n/a	n/a	n/a
66	1833937	3619688H1	SNP00032201	238	457	C	C	T	noncoding	n/a	n/a	n/a	n/a
66	1833937	3619688H1	SNP00142786	83	302	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	1833937	389627H1	SNP00032200	163	195	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	1833937	3994121H1	SNP00032201	231	448	C	C	T	noncoding	n/a	n/a	n/a	n/a
66	1833937	3994121H1	SNP00142786	76	293	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	1833937	3996413H1	SNP00142786	83	301	A	A	G	noncoding	n/a	n/a	n/a	n/a
66	1833937	4504232H1	SNP00119902	143	167	T	T	C	S49	n/a	n/a	n/a	n/a
66	1833937	4528504H1	SNP00043506	76	82	G	G	T	K20	n/d	n/a	n/a	n/a
66	1833937	5490843H1	SNP00051244	145	143	T	T	G	F41	n/a	n/a	n/a	n/a
66	1833937	6384674H1	SNP00051244	169	160	T	T	G	N46	n/a	n/a	n/a	n/a
66	1833937	6558901H1	SNP00119902	277	172	T	T	C	T50	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
67	7502036	1639008H1	SNP00021105	23	1547	T	C	T	H435	n/d	n/a	n/a	n/a
67	7502036	2534906H2	SNP00113937	42	499	A	A	G	Q86	n/a	n/a	n/a	n/a
67	7502036	2534906H2	SNP00116337	165	622	T	T	G	V127	0.97	0.93	0.94	0.96
67	7502036	2534906H2	SNP00135699	143	600	G	G	T	V120	n/a	n/a	n/a	n/a
67	7502036	3074078H1	SNP00002808	163	198	G	A	G	noncoding	n/a	n/a	n/a	n/a
67	7502036	3646156H1	SNP00113937	183	497	A	A	G	E85	n/a	n/a	n/a	n/a
67	7502036	4786529H1	SNP00135699	126	579	G	G	T	V113	n/a	n/a	n/a	n/a
67	7502036	4889178H1	SNP00116337	196	620	T	T	G	L126	0.97	0.93	0.94	0.96
67	7502036	4889178H1	SNP00135699	174	598	G	G	T	G119	n/a	n/a	n/a	n/a
67	7502036	5063984H1	SNP00116336	118	370	A	A	G	N43	n/d	n/d	n/d	n/d
67	7502036	5117894H1	SNP00002808	160	197	A	A	G	noncoding	n/a	n/a	n/a	n/a
67	7502036	6445802H2	SNP00002808	122	186	G	A	G	noncoding	n/a	n/a	n/a	n/a
67	7502036	7112838H2	SNP00120168	408	784	T	C	T	V181	0.89	0.93	0.94	0.91
68	7503248	5952116H1	SNP00122931	240	385	A	A	C	G88	n/a	n/a	n/a	n/a
68	7503248	6917805H1	SNP00122931	48	381	A	A	C	Q87	n/a	n/a	n/a	n/a
68	7503248	7136072H1	SNP00116205	129	1089	G	G	A	noncoding	n/a	n/a	n/a	n/a
70	7505931	3734834H1	SNP00043514	60	540	C	C	A	noncoding	n/d	n/a	n/a	n/a
70	7505931	5873534H1	SNP00115130	263	262	G	G	A	E54	n/a	n/a	n/a	n/a
70	7505931	6563752H1	SNP00115130	264	265	G	G	A	Q55	n/a	n/a	n/a	n/a
71	7506912	2913521H1	SNP00043631	155	1172	C	C	T	H387	0.98	n/a	n/a	n/a
71	7506912	3281384H1	SNP00043630	89	980	T	T	C	C323	n/a	n/a	n/a	n/a
71	7506912	3463658H1	SNP00043630	57	979	T	T	C	I322	n/a	n/a	n/a	n/a
71	7506912	3907226H1	SNP00043630	120	982	T	T	C	C323	n/a	n/a	n/a	n/a
71	7506912	4286341H1	SNP00043629	78	319	T	C	T	F102	n/a	n/a	n/a	n/a
71	7506912	4443694H1	SNP00043630	155	978	C	T	C	T322	n/a	n/a	n/a	n/a
71	7506912	7363877H1	SNP00043631	103	1174	C	C	T	D387	0.98	n/a	n/a	n/a
71	7506912	7637420H1	SNP00043629	484	321	C	C	T	P103	n/a	n/a	n/a	n/a
72	7506913	2913521H1	SNP00043631	155	937	C	C	T	H313	0.98	n/a	n/a	n/a
72	7506913	3281384H1	SNP00043630	89	745	T	T	C	C249	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
72	7506913	3463658H1	SNP00043630	57	744	T	T	C	I248	n/a	n/a	n/a	n/a
72	7506913	3907226H1	SNP00043630	120	747	T	T	C	C249	n/a	n/a	n/a	n/a
72	7506913	4286341H1	SNP00043629	78	261	T	C	T	F87	n/a	n/a	n/a	n/a
72	7506913	4443694H1	SNP00043630	155	743	C	T	C	T248	n/a	n/a	n/a	n/a
72	7506913	7363877H1	SNP00043631	103	939	C	C	T	D313	0.98	n/a	n/a	n/a
72	7506913	7637420H1	SNP00043629	484	263	C	C	T	P88	n/a	n/a	n/a	n/a
74	7507063	2984716H1	SNP00036581	27	723	C	T	C	noncoding	n/d	n/a	n/a	n/a
74	7507063	7395779H1	SNP00036581	58	717	T	T	C	noncoding	n/d	n/a	n/a	n/a
76	7509265	1824201H1	SNP00066979	79	1120	C	C	T	noncoding	n/d	n/d	n/a	n/d
76	7509265	2046231H1	SNP00114113	228	1399	G	G	C	noncoding	n/d	n/a	n/a	n/a
76	7509265	2744627H1	SNP00022802	98	167	C	C	G	noncoding	n/a	n/a	n/a	n/a
76	7509265	3054645H1	SNP00066979	68	1119	C	C	T	noncoding	n/d	n/d	n/a	n/d
76	7509265	3492810H1	SNP00022802	97	165	G	C	G	noncoding	n/a	n/a	n/a	n/a
76	7509265	4146813H1	SNP00149600	241	2045	T	T	C	noncoding	n/a	n/a	n/a	n/a
76	7509265	5059890H1	SNP00114113	10	1397	G	G	C	noncoding	n/d	n/a	n/a	n/a
76	7509265	6200476H1	SNP00114113	124	1394	G	G	C	noncoding	n/d	n/a	n/a	n/a
76	7509265	6200676H1	SNP00114113	129	1400	G	G	C	noncoding	n/d	n/a	n/a	n/a
76	7509265	6537489H1	SNP00149600	174	2043	T	T	C	noncoding	n/a	n/a	n/a	n/a
76	7509265	7622751H1	SNP00031991	177	1165	C	C	T	noncoding	n/a	n/a	n/a	n/a
77	7509371	7990470H2	SNP00135214	206	333	G	A	G	L81	n/a	n/a	n/a	n/a
82	7509246	2984716H1	SNP00036581	27	723	C	T	C	noncoding	n/d	n/a	n/a	n/a
82	7509246	7395779H1	SNP00036581	58	717	T	T	C	noncoding	n/d	n/a	n/a	n/a

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to the amino acid sequence of SEQ ID NO:1-2, SEQ ID NO:6, SEQ ID NO:9, SEQ ID NO:14, SEQ ID NO:16-18, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:34-38, and SEQ ID NO:40-42,

c) a polypeptide comprising a naturally occurring amino acid sequence at least 92% identical to the amino acid sequence of SEQ ID NO:24,

d) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:19, SEQ ID NO:29, SEQ ID NO:32-33,

e) a polypeptide comprising a naturally occurring amino acid sequence at least 96% identical to the amino acid sequence of SEQ ID NO:13,

f) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:7, SEQ ID NO:11 and SEQ ID NO:26,

g) a polypeptide comprising a naturally occurring amino acid sequence at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:20 and SEQ ID NO:30,

h) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:3-5, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:15, SEQ ID NO:21, SEQ ID NO:28 and SEQ ID NO:31,

i) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and

j) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the

group consisting of SEQ ID NO:1-42.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least

90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-61, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71-74, and SEQ ID NO:77-84,

c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 91% identical to the polynucleotide sequence of SEQ ID NO:68,

d) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 93% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:63-64 and SEQ ID NO:75,

e) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 95% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:62 and SEQ ID NO:65,

f) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 96% identical to the polynucleotide sequence of SEQ ID NO:76,

g) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 98% identical to the polynucleotide sequence of SEQ ID NO:66,

h) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to the amino acid sequence of SEQ ID NO:70,

i) a polynucleotide complementary to a polynucleotide of a),

j) a polynucleotide complementary to a polynucleotide of b),

k) a polynucleotide complementary to a polynucleotide of c),

l) a polynucleotide complementary to a polynucleotide of d),

m) a polynucleotide complementary to a polynucleotide of e),

n) a polynucleotide complementary to a polynucleotide of f),

o) a polynucleotide complementary to a polynucleotide of g),

p) a polynucleotide complementary to a polynucleotide of h), and

q) an RNA equivalent of a)-p).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

19. A method for treating a disease or condition associated with decreased expression of functional KPP, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional KPP, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional KPP, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
c) comparing the activity of the polypeptide of claim 1 in the presence of the test

compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

a) treating a biological sample containing nucleic acids with the test compound,

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A method for a diagnostic test for a condition or disease associated with the expression of KPP in a biological sample, the method comprising:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of KPP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, further comprising a label.

35. A method of diagnosing a condition or disease associated with the expression of KPP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from the animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibody producing cells from the animal,

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,

d) culturing the hybridoma cells, and

e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42 in a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42 from a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:43.

99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:44.

100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:45.

101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:46.

102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:47.

103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:48.

104. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:49.

105. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:50.

106. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:51.

107. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:52.

108. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:53.

109. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:54.

110. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:55.

111. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:56.

112. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:57.

113. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:58.

114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:59.

115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:60.

116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:61.

117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:63.

119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:64.

120. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:65.

121. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:66.

122. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:84.

```
<110> INCYTE GENOMICS, INC.
      KABLE, Amy E.
      CHIEN, David
      WILSON, Amy D.
      SWARNAKAR, Anita
      GORVAD, Ann E.
      HAFALIA, April J.A.
      EMERLING, Brooke M.
      RAMKUMAR, Jayalaxmi
      JIN, Pei
      GRIFFIN, Jennifer A.
      MARQUIS, Joseph P.
      BAUGHN, Mariah R.
      CHAWLA, Narinder K.
      LEHR-MASON, Patricia M.
      KHARE, Reena
      LEE, Sally
      HAWKINS, Phillip R.
      BECHA, Shanya D.
      LEE, Soo Yeun
      SPRAGUE, William W.
      ZEBARJADIAN, Yeganeh

<120> KINASES AND PHOSPHATASES

<130> PF-1315 PCT

<140> To Be Assigned
<141> Herewith

<150> US 60/340,235
<151> 2001-12-07

<150> US 60/343,007
<151> 2001-12-19

<150> US 60/343,546
<151> 2001-12-21

<150> US 60/354,388
<151> 2002-02-04

<150> US 60/357,675
<151> 2002-02-15

<160> 84
<170> PERL Program

<210> 1
<211> 1179
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503679CD1
```

```
<400> 1
Met His Val Phe Ile Val Gly Met Ser Leu Ser Ser Val Thr Leu
 1               5                  10                  15
Ala Ser Ala Leu Gln Val Arg Gly Glu Ala Leu Ser Glu Glu Glu
                20                  25                  30
Ile Trp Ser Leu Leu Phe Leu Ala Ala Glu Gln Leu Leu Glu Asp
                35                  40                  45
Leu Arg Asn Asp Ser Ser Asp Tyr Val Val Cys Pro Trp Ser Ala
                50                  55                  60
Leu Leu Ser Ala Ala Gly Ser Leu Ser Phe Gln Gly Arg Val Ser
                65                  70                  75
His Ile Glu Ala Ala Pro Phe Lys Ala Pro Glu Leu Leu Gln Gly
                80                  85                  90
Gln Ser Glu Asp Glu Gln Pro Asp Ala Ser Gln Pro Leu Gln Leu
                95                 100                 105
Cys Glu Pro Leu His Ser Ile Leu Leu Thr Met Cys Glu Asp Gln
               110                 115                 120
Pro His Arg Arg Cys Thr Leu Gln Ser Val Leu Glu Ala Cys Arg
               125                 130                 135
Val His Glu Lys Glu Val Ser Val Tyr Pro Ala Pro Ala Gly Leu
               140                 145                 150
His Ile Arg Arg Leu Val Gly Leu Val Leu Gly Thr Ile Ser Glu
               155                 160                 165
Val Glu Lys Arg Val Val Glu Glu Ser Ser Ser Val Gln Gln Asn
               170                 175                 180
Arg Ser Tyr Leu Leu Arg Lys Arg Leu Arg Gly Thr Ser Ser Glu
               185                 190                 195
Ser Pro Ala Ala Gln Ala Pro Glu Cys Leu His Pro Cys Arg Val
               200                 205                 210
Ser Glu Arg Ser Thr Glu Thr Gln Ser Ser Pro Glu Pro His Trp
               215                 220                 225
Ser Thr Leu Thr His Ser His Cys Ser Leu Leu Val Asn Arg Ala
               230                 235                 240
Leu Pro Gly Ala Asp Pro Gln Asp Gln Gln Ala Gly Arg Arg Leu
               245                 250                 255
Ser Ser Gly Ser Val His Ser Ala Ala Asp Ser Ser Trp Pro Thr
               260                 265                 270
Thr Pro Ser Gln Arg Gly Phe Leu Gln Arg Arg Ser Lys Phe Ser
               275                 280                 285
Arg Pro Glu Phe Ile Leu Leu Ala Gly Glu Ala Pro Met Thr Leu
               290                 295                 300
His Leu Pro Gly Ser Val Val Thr Lys Lys Gly Lys Ser Tyr Leu
               305                 310                 315
Ala Leu Arg Asp Leu Cys Val Val Leu Leu Asn Gly Gln His Leu
               320                 325                 330
Glu Val Lys Cys Asp Val Glu Ser Thr Val Gly Ala Val Phe Asn
               335                 340                 345
Ala Val Thr Ser Phe Ala Asn Leu Glu Glu Leu Thr Tyr Phe Gly
               350                 355                 360
Leu Ala Tyr Met Lys Ser Lys Glu Phe Phe Phe Leu Asp Ser Glu
               365                 370                 375
Thr Arg Leu Cys Lys Ile Ala Pro Glu Gly Trp Arg Glu Gln Pro
               380                 385                 390
Gln Lys Thr Ser Met Asn Thr Phe Thr Leu Phe Leu Arg Ile Lys
               395                 400                 405
Phe Phe Val Ser His Tyr Gly Leu Leu Gln His Ser Leu Thr Arg
```

```
                410                 415                 420
His Gln Phe Tyr Leu Gln Leu Arg Lys Asp Ile Leu Glu Glu Arg
                425                 430                 435
Leu Tyr Cys Asn Glu Glu Ile Leu Leu Gln Leu Gly Val Leu Ala
                440                 445                 450
Leu Gln Ala Glu Phe Gly Asn Tyr Pro Lys Glu Gln Val Glu Ser
                455                 460                 465
Lys Pro Tyr Phe His Val Glu Asp Tyr Ile Pro Ala Ser Leu Ile
                470                 475                 480
Glu Arg Met Thr Ala Leu Arg Val Gln Val Glu Val Ser Glu Met
                485                 490                 495
His Arg Leu Ser Ser Ala Leu Trp Gly Glu Asp Ala Glu Leu Lys
                500                 505                 510
Phe Leu Arg Val Thr Gln Gln Leu Pro Glu Tyr Gly Val Leu Val
                515                 520                 525
His Gln Val Phe Ser Glu Lys Arg Arg Pro Glu Glu Glu Met Ala
                530                 535                 540
Leu Gly Ile Cys Ala Lys Gly Val Ile Val Tyr Glu Val Lys Asn
                545                 550                 555
Asn Ser Arg Ile Ala Met Leu Arg Phe Gln Trp Arg Glu Thr Gly
                560                 565                 570
Lys Ile Ser Thr Tyr Gln Lys Lys Phe Thr Ile Thr Ser Ser Val
                575                 580                 585
Thr Gly Lys Lys His Thr Phe Val Thr Asp Ser Ala Lys Thr Ser
                590                 595                 600
Lys Tyr Leu Leu Asp Leu Cys Ser Ala Gln His Gly Phe Asn Ala
                605                 610                 615
Gln Met Gly Ser Gly Gln Pro Ser His Val Leu Phe Asp His Asp
                620                 625                 630
Lys Phe Val Gln Met Ala Asn Leu Ser Pro Ala His Gln Ala Arg
                635                 640                 645
Ser Lys Pro Leu Ile Trp Ile Gln Arg Leu Ser Cys Ser Glu Asn
                650                 655                 660
Glu Leu Phe Val Ser Arg Leu Gln Gly Ala Ala Gly Gly Leu Leu
                665                 670                 675
Ser Thr Ser Met Asp Asn Phe Asn Val Asp Gly Ser Lys Glu Ala
                680                 685                 690
Gly Ala Glu Gly Ile Gly Arg Ser Pro Cys Thr Gly Arg Glu Gln
                695                 700                 705
Leu Lys Ser Ala Cys Val Ile Gln Lys Pro Met Thr Trp Asp Ser
                710                 715                 720
Leu Ser Gly Pro Pro Val Gln Ser Met His Ala Gly Val Gly Gly
                725                 730                 735
Asn Asn Pro Asp Glu Glu Lys Asn Gly Thr Ala Asn Ser Gly Val
                740                 745                 750
Ser Ser Thr Asp Ile Leu Ser Phe Gly Tyr Gln Gly Ser Leu Leu
                755                 760                 765
Ser His Thr Gln Asp Gln Asp Arg Asn Thr Glu Glu Leu Asp Met
                770                 775                 780
Ala Gly Val Gln Ser Leu Val Pro Arg Leu Arg His Gln Leu Ser
                785                 790                 795
Phe Leu Pro Leu Lys Gly Ala Gly Ser Ser Cys Pro Pro Ser Pro
                800                 805                 810
Pro Glu Ile Ser Ala Gly Glu Ile Tyr Phe Val Glu Leu Val Lys
                815                 820                 825
Glu Asp Gly Thr Leu Gly Phe Ser Val Thr Gly Gly Ile Asn Thr
```

```
                    830                     835                     840
Ser Val Pro Tyr Gly Gly Ile Tyr Val Lys Ser Ile Val Pro Gly
                    845                     850                     855
Gly Pro Ala Ala Lys Glu Gly Gln Ile Leu Gln Gly Asp Arg Leu
                    860                     865                     870
Leu Gln Val Asp Gly Val Ile Leu Cys Gly Leu Thr His Lys Gln
                    875                     880                     885
Ala Val Gln Cys Leu Lys Gly Pro Gly Gln Val Ala Arg Leu Val
                    890                     895                     900
Leu Glu Arg Arg Val Pro Arg Ser Thr Gln Gln Cys Pro Ser Ala
                    905                     910                     915
Asn Asp Ser Met Gly Asp Glu Arg Thr Ala Val Ser Leu Val Thr
                    920                     925                     930
Ala Leu Pro Gly Arg Pro Ser Ser Cys Val Ser Val Thr Asp Gly
                    935                     940                     945
Pro Lys Phe Glu Val Lys Leu Lys Lys Asn Ala Asn Gly Leu Gly
                    950                     955                     960
Phe Ser Phe Val Gln Met Glu Lys Glu Ser Cys Ser His Leu Lys
                    965                     970                     975
Ser Asp Leu Val Arg Ile Lys Arg Leu Phe Pro Gly Gln Pro Ala
                    980                     985                     990
Glu Glu Asn Gly Ala Ile Ala Ala Gly Asp Ile Ile Leu Ala Val
                    995                    1000                    1005
Asn Gly Arg Ser Thr Glu Gly Leu Ile Phe Gln Glu Val Leu His
                   1010                    1015                    1020
Leu Leu Arg Gly Ala Pro Gln Glu Val Thr Leu Leu Leu Cys Arg
                   1025                    1030                    1035
Pro Pro Pro Gly Ala Leu Pro Glu Leu Glu Gln Glu Trp Gln Thr
                   1040                    1045                    1050
Pro Glu Leu Ser Ala Asp Lys Glu Phe Thr Arg Ala Thr Cys Thr
                   1055                    1060                    1065
Asp Ser Cys Thr Ser Pro Ile Leu Asp Gln Glu Asp Ser Trp Arg
                   1070                    1075                    1080
Asp Ser Ala Ser Pro Asp Ala Gly Glu Gly Leu Gly Leu Arg Pro
                   1085                    1090                    1095
Glu Ser Ser Gln Lys Ala Ile Arg Glu Ala Gln Trp Gly Gln Asn
                   1100                    1105                    1110
Arg Glu Arg Pro Trp Ala Ser Ser Leu Thr His Ser Pro Glu Ser
                   1115                    1120                    1125
His Pro His Leu Cys Lys Leu His Gln Glu Arg Asp Glu Ser Thr
                   1130                    1135                    1140
Leu Ala Thr Ser Leu Glu Lys Asp Val Arg Gln Asn Cys Tyr Ser
                   1145                    1150                    1155
Val Cys Asp Ile Met Arg Leu Gly Arg Tyr Ser Phe Ser Ser Pro
                   1160                    1165                    1170
Leu Thr Arg Leu Ser Thr Asp Ile Phe
                   1175

<210> 2
<211> 1170
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503681CD1
```

```
<400> 2
Met Cys Glu Asp Gln Pro His Arg Arg Cys Thr Leu Gln Ser Val
 1              5                   10                  15
Leu Glu Ala Cys Arg Val His Glu Lys Glu Val Ser Val Tyr Pro
                20                  25                  30
Ala Pro Ala Gly Leu His Ile Arg Arg Leu Val Gly Leu Val Leu
                35                  40                  45
Gly Thr Ile Ser Glu Val Glu Lys Arg Val Val Glu Glu Ser Ser
                50                  55                  60
Ser Val Gln Gln Asn Arg Ser Tyr Leu Leu Arg Lys Arg Leu Arg
                65                  70                  75
Gly Thr Ser Ser Glu Ser Pro Ala Ala Gln Ala Pro Glu Cys Leu
                80                  85                  90
His Pro Cys Arg Val Ser Glu Arg Ser Thr Glu Thr Gln Ser Ser
                95                  100                 105
Pro Glu Pro His Trp Ser Thr Leu Thr His Ser His Cys Ser Leu
                110                 115                 120
Leu Val Asn Arg Ala Leu Pro Gly Ala Asp Pro Gln Asp Gln Gln
                125                 130                 135
Ala Gly Arg Arg Leu Ser Ser Gly Ser Val His Ser Ala Ala Asp
                140                 145                 150
Ser Ser Trp Pro Thr Thr Pro Ser Gln Arg Gly Phe Leu Gln Arg
                155                 160                 165
Arg Ser Lys Phe Ser Arg Pro Glu Phe Ile Leu Leu Ala Gly Glu
                170                 175                 180
Ala Pro Met Thr Leu His Leu Pro Gly Ser Val Val Thr Lys Lys
                185                 190                 195
Gly Lys Ser Tyr Leu Ala Leu Arg Asp Leu Cys Val Val Leu Leu
                200                 205                 210
Asn Gly Gln His Leu Glu Val Lys Cys Asp Val Glu Ser Thr Val
                215                 220                 225
Gly Ala Val Phe Asn Ala Val Thr Ser Phe Ala Asn Leu Glu Glu
                230                 235                 240
Leu Thr Tyr Phe Gly Leu Ala Tyr Met Lys Ser Lys Glu Phe Phe
                245                 250                 255
Phe Leu Asp Ser Glu Thr Arg Leu Cys Lys Ile Ala Pro Glu Gly
                260                 265                 270
Trp Arg Glu Gln Pro Gln Lys Thr Ser Met Asn Thr Phe Thr Leu
                275                 280                 285
Phe Leu Arg Ile Lys Phe Phe Val Ser His Tyr Gly Leu Leu Gln
                290                 295                 300
His Ser Leu Thr Arg His Gln Phe Tyr Leu Gln Leu Arg Lys Asp
                305                 310                 315
Ile Leu Glu Glu Arg Leu Tyr Cys Asn Glu Glu Ile Leu Leu Gln
                320                 325                 330
Leu Gly Val Leu Ala Leu Gln Ala Glu Phe Gly Asn Tyr Pro Lys
                335                 340                 345
Glu Gln Val Glu Ser Lys Pro Tyr Phe His Val Glu Asp Tyr Ile
                350                 355                 360
Pro Ala Ser Leu Ile Glu Arg Met Thr Ala Leu Arg Val Gln Val
                365                 370                 375
Glu Val Ser Glu Met His Arg Leu Ser Ser Ala Leu Trp Gly Glu
                380                 385                 390
Asp Ala Glu Leu Lys Phe Leu Arg Val Thr Gln Gln Leu Pro Glu
                395                 400                 405
Tyr Gly Val Leu Val His Gln Val Phe Ser Glu Lys Arg Arg Pro
```

```
                410                 415                 420
Glu Glu Glu Met Ala Leu Gly Ile Cys Ala Lys Gly Val Ile Val
                425                 430                 435
Tyr Glu Val Lys Asn Asn Ser Arg Ile Ala Met Leu Arg Phe Gln
                440                 445                 450
Trp Arg Glu Thr Gly Lys Ile Ser Thr Tyr Gln Lys Lys Phe Thr
                455                 460                 465
Ile Thr Ser Ser Val Thr Gly Lys Lys His Thr Phe Val Thr Asp
                470                 475                 480
Ser Ala Lys Thr Ser Lys Tyr Leu Leu Asp Leu Cys Ser Ala Gln
                485                 490                 495
His Gly Phe Asn Ala Gln Met Gly Ser Gly Gln Pro Ser His Val
                500                 505                 510
Leu Phe Asp His Asp Lys Phe Val Gln Met Ala Asn Leu Ser Pro
                515                 520                 525
Ala His Gln Ala Arg Ser Lys Pro Leu Ile Trp Ile Gln Arg Leu
                530                 535                 540
Ser Cys Ser Glu Asn Glu Leu Phe Val Ser Arg Leu Gln Gly Ala
                545                 550                 555
Ala Gly Gly Leu Leu Ser Thr Ser Met Asp Asn Phe Asn Val Asp
                560                 565                 570
Gly Ser Lys Glu Ala Gly Ala Glu Gly Ile Gly Arg Ser Pro Cys
                575                 580                 585
Thr Gly Arg Glu Gln Leu Lys Ser Ala Cys Val Ile Gln Lys Pro
                590                 595                 600
Met Thr Trp Asp Ser Leu Ser Gly Pro Pro Val Gln Ser Met His
                605                 610                 615
Ala Gly Ser Lys Asn Asn Arg Arg Lys Ser Phe Ile Ala Glu Pro
                620                 625                 630
Gly Arg Glu Ile Val Arg Val Thr Leu Lys Arg Asp Pro His Arg
                635                 640                 645
Gly Phe Gly Phe Val Ile Asn Glu Gly Glu Tyr Ser Gly Gln Ala
                650                 655                 660
Asp Pro Gly Ile Phe Ile Ser Ser Ile Ile Pro Gly Gly Pro Ala
                665                 670                 675
Glu Lys Ala Lys Thr Ile Lys Pro Gly Gly Gln Ile Leu Ala Leu
                680                 685                 690
Asn His Ile Ser Leu Glu Gly Phe Thr Phe Asn Met Ala Val Arg
                695                 700                 705
Met Ile Gln Asn Ser Pro Asp Asn Ile Glu Leu Ile Ile Ser Gln
                710                 715                 720
Ser Lys Gly Val Gly Gly Asn Asn Pro Asp Glu Glu Lys Asn Ser
                725                 730                 735
Thr Ala Asn Ser Gly Val Ser Ser Thr Asp Val Leu Ser Phe Gly
                740                 745                 750
Tyr Gln Gly Ser Leu Leu Ser His Thr Gln Asp Gln Asp Arg Asn
                755                 760                 765
Thr Glu Glu Leu Asp Met Ala Gly Val Gln Ser Leu Val Pro Arg
                770                 775                 780
Leu Arg His Gln Leu Ser Phe Leu Pro Leu Lys Gly Ala Gly Ser
                785                 790                 795
Ser Cys Pro Pro Ser Pro Pro Glu Ile Ser Ala Gly Glu Ile Tyr
                800                 805                 810
Phe Val Glu Leu Val Lys Glu Asp Gly Thr Leu Gly Phe Ser Val
                815                 820                 825
Thr Gly Gly Ile Asn Thr Ser Val Pro Tyr Gly Gly Ile Tyr Val
```

```
                830                 835                 840
Lys Ser Ile Val Pro Gly Gly Pro Ala Ala Lys Glu Gly Gln Ile
                845                 850                 855
Leu Gln Gly Asp Arg Leu Leu Gln Val Asp Gly Val Ile Leu Cys
                860                 865                 870
Gly Leu Thr His Lys Gln Ala Val Gln Cys Leu Lys Gly Pro Gly
                875                 880                 885
Gln Val Ala Arg Leu Val Leu Glu Arg Arg Val Pro Arg Ser Thr
                890                 895                 900
Gln Gln Cys Pro Ser Ala Asn Asp Ser Met Gly Asp Glu Arg Thr
                905                 910                 915
Ala Val Ser Leu Val Thr Ala Leu Pro Gly Arg Pro Ser Ser Cys
                920                 925                 930
Val Ser Val Thr Asp Gly Pro Lys Phe Glu Val Lys Leu Lys Lys
                935                 940                 945
Asn Ala Asn Gly Leu Gly Phe Ser Phe Val Gln Met Glu Lys Glu
                950                 955                 960
Ser Cys Ser His Leu Lys Ser Asp Leu Val Arg Ile Lys Arg Leu
                965                 970                 975
Phe Pro Gly Gln Pro Ala Glu Glu Asn Gly Ala Ile Ala Ala Gly
                980                 985                 990
Asp Ile Ile Leu Ala Val Asn Gly Arg Ser Thr Glu Gly Leu Ile
                995                1000                1005
Phe Gln Glu Val Leu His Leu Leu Arg Gly Ala Pro Gln Glu Val
               1010                1015                1020
Thr Leu Leu Leu Cys Arg Pro Pro Pro Gly Ala Leu Pro Glu Leu
               1025                1030                1035
Glu Gln Glu Trp Gln Thr Pro Glu Leu Ser Ala Asp Lys Glu Phe
               1040                1045                1050
Thr Arg Ala Thr Cys Thr Asp Ser Cys Thr Ser Pro Ile Leu Asp
               1055                1060                1065
Gln Glu Asp Ser Trp Arg Asp Ser Ala Ser Pro Asp Ala Gly Glu
               1070                1075                1080
Gly Leu Gly Leu Arg Pro Glu Ser Ser Gln Lys Ala Ile Arg Glu
               1085                1090                1095
Ala Gln Trp Gly Gln Asn Arg Glu Arg Pro Trp Ala Ser Ser Leu
               1100                1105                1110
Thr His Ser Pro Glu Ser His Pro His Leu Cys Lys Leu His Gln
               1115                1120                1125
Glu Arg Asp Glu Ser Thr Leu Ala Thr Ser Leu Glu Lys Asp Val
               1130                1135                1140
Arg Gln Asn Cys Tyr Ser Val Cys Asp Ile Met Arg Leu Gly Arg
               1145                1150                1155
Tyr Ser Phe Ser Ser Pro Leu Thr Arg Leu Ser Thr Asp Ile Phe
               1160                1165                1170


<210> 3
<211> 1977
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505819CD1
```

```
<400> 3
Met Phe Leu Lys Cys Asp Met Ser Ile His Asp Val Ser Val Ser
  1               5                  10                  15
Ala Ala Glu Thr Gln Ala Ile Val Asp Glu Ala Leu Gly Leu Arg
                 20                  25                  30
Lys Lys Arg Gln Ala Leu Ile Val Arg Glu Lys Glu Pro Asp Leu
                 35                  40                  45
Lys Leu Val Gln Pro Ile Pro Phe Phe Thr Trp Lys Cys Leu Gly
                 50                  55                  60
Glu Ser Leu Leu Ala Met Tyr Asn His Leu Thr Thr Cys Glu Pro
                 65                  70                  75
Pro Arg Pro Ser Leu Gly Lys Arg Ile Asp Leu Ser Asp Tyr Gln
                 80                  85                  90
Asp Pro Ser Gln Pro Leu Glu Ser Ser Met Val Val Thr Pro Val
                 95                 100                 105
Asn Val Ile Gln Pro Ser Thr Val Ser Thr Asn Pro Ala Val Ala
                110                 115                 120
Val Ala Glu Pro Val Val Ser Tyr Thr Ser Val Ala Thr Thr Ser
                125                 130                 135
Phe Pro Leu His Ser Pro Gly Leu Leu Glu Thr Gly Ala Pro Val
                140                 145                 150
Gly Asp Ile Ser Gly Gly Asp Lys Ser Lys Lys Gly Val Lys Arg
                155                 160                 165
Lys Lys Ile Ser Glu Glu Ser Gly Glu Thr Ala Lys Arg Arg Ser
                170                 175                 180
Ala Arg Val Arg Asn Thr Lys Cys Lys Lys Glu Glu Lys Val Asp
                185                 190                 195
Phe Gln Glu Leu Leu Met Lys Phe Leu Pro Ser Arg Leu Arg Lys
                200                 205                 210
Leu Asp Pro Glu Glu Glu Asp Asp Ser Phe Asn Asn Tyr Glu Val
                215                 220                 225
Gln Ser Glu Ala Lys Leu Glu Ser Phe Pro Ser Ile Gly Pro Gln
                230                 235                 240
Arg Leu Ser Phe Asp Ser Ala Thr Phe Met Glu Ser Glu Lys Gln
                245                 250                 255
Asp Val His Glu Phe Leu Leu Glu Asn Leu Thr Asn Gly Gly Ile
                260                 265                 270
Leu Glu Leu Met Met Arg Tyr Leu Lys Ala Met Gly His Lys Phe
                275                 280                 285
Leu Val Arg Trp Pro Pro Gly Leu Ala Glu Val Val Leu Ser Val
                290                 295                 300
Tyr His Ser Trp Arg Arg His Ser Thr Ser Leu Pro Asn Pro Leu
                305                 310                 315
Leu Arg Asp Cys Ser Asn Lys His Ile Lys Asp Arg Met Leu Met
                320                 325                 330
Ser Leu Ser Cys Met Glu Leu Gln Leu Asp Gln Trp Leu Leu Thr
                335                 340                 345
Lys Gly Arg Ser Ser Ala Val Ser Pro Arg Asn Cys Pro Ala Gly
                350                 355                 360
Met Val Asn Gly Arg Phe Gly Pro Asp Phe Pro Gly Thr His Cys
                365                 370                 375
Leu Gly Asp Leu Leu Gln Leu Ser Phe Ala Ser Ser Gln Arg Asp
                380                 385                 390
Leu Phe Glu Asp Gly Trp Leu Glu Phe Val Val Arg Val Tyr Trp
                395                 400                 405
Leu Lys Ala Arg Phe Leu Ala Leu Gln Gly Asp Met Glu Gln Ala
```

```
                    410                 415                 420
Leu Glu Asn Tyr Asp Ile Cys Thr Glu Met Leu Gln Ser Ser Thr
                    425                 430                 435
Ala Ile Gln Val Glu Ala Gly Ala Glu Arg Arg Asp Ile Val Ile
                    440                 445                 450
Arg Leu Pro Asn Leu His Asn Asp Ser Val Val Ser Leu Glu Glu
                    455                 460                 465
Ile Asp Lys Asn Leu Lys Ser Leu Glu Arg Cys Gln Ser Leu Glu
                    470                 475                 480
Glu Ile Gln Arg Leu Tyr Glu Ala Gly Asp Tyr Lys Ala Val Val
                    485                 490                 495
His Leu Leu Arg Pro Thr Leu Cys Thr Ser Gly Phe Asp Arg Ala
                    500                 505                 510
Lys His Leu Glu Phe Met Thr Ser Ile Pro Glu Arg Pro Ala Gln
                    515                 520                 525
Leu Leu Leu Leu Gln Asp Ser Leu Leu Arg Leu Lys Asp Tyr Arg
                    530                 535                 540
Gln Cys Phe Glu Cys Ser Asp Val Ala Leu Asn Glu Ala Val Gln
                    545                 550                 555
Gln Met Val Asn Ser Gly Glu Ala Ala Ala Lys Glu Glu Trp Val
                    560                 565                 570
Ala Thr Val Thr Gln Leu Leu Met Gly Ile Glu Gln Ala Leu Ser
                    575                 580                 585
Ala Asp Ser Ser Gly Ser Ile Leu Lys Val Ser Ser Ser Thr Thr
                    590                 595                 600
Gly Leu Val Arg Leu Thr Asn Asn Leu Ile Gln Val Ile Asp Cys
                    605                 610                 615
Ser Met Ala Val Gln Glu Glu Ala Lys Glu Pro His Val Ser Ser
                    620                 625                 630
Val Leu Pro Trp Ile Ile Leu His Arg Ile Ile Trp Gln Glu Glu
                    635                 640                 645
Asp Thr Phe His Ser Leu Cys His Gln Gln Gln Leu Gln Asn Pro
                    650                 655                 660
Ala Glu Glu Gly Met Ser Glu Thr Pro Met Leu Pro Ser Ser Leu
                    665                 670                 675
Met Leu Leu Asn Thr Ala His Glu Tyr Leu Gly Arg Arg Ser Trp
                    680                 685                 690
Cys Cys Asn Ser Asp Gly Ala Leu Leu Arg Phe Tyr Val Arg Val
                    695                 700                 705
Leu Gln Lys Glu Leu Ala Ala Ser Thr Ser Glu Asp Thr His Pro
                    710                 715                 720
Tyr Lys Glu Glu Leu Glu Thr Ala Leu Glu Gln Cys Phe Tyr Cys
                    725                 730                 735
Leu Tyr Ser Phe Pro Ser Lys Lys Ser Lys Ala Arg Tyr Leu Glu
                    740                 745                 750
Glu His Ser Ala Gln Gln Val Asp Leu Ile Trp Glu Asp Ala Leu
                    755                 760                 765
Phe Met Phe Glu Tyr Phe Lys Pro Lys Thr Leu Pro Glu Phe Asp
                    770                 775                 780
Ser Tyr Lys Thr Ser Thr Val Ser Ala Asp Leu Ala Asn Leu Leu
                    785                 790                 795
Lys Arg Ile Ala Thr Ile Val Pro Arg Thr Glu Arg Pro Ala Leu
                    800                 805                 810
Ser Leu Asp Lys Val Ser Ala Tyr Ile Glu Gly Thr Ser Thr Glu
                    815                 820                 825
Val Pro Cys Leu Pro Glu Gly Ala Asp Pro Ser Pro Pro Val Val
```

```
                830                 835                 840
Asn Glu Leu Tyr Tyr Leu Leu Ala Asp Tyr His Phe Lys Asn Lys
                845                 850                 855
Glu Gln Ser Lys Ala Ile Lys Phe Tyr Met His Asp Ile Cys Ile
                860                 865                 870
Cys Pro Asn Arg Phe Asp Ser Trp Ala Gly Met Ala Leu Ala Arg
                875                 880                 885
Ala Ser Arg Ile Gln Asp Lys Leu Asn Ser Asn Glu Leu Lys Ser
                890                 895                 900
Asp Gly Pro Ile Trp Lys His Ala Thr Pro Val Leu Asn Cys Phe
                905                 910                 915
Arg Arg Ala Leu Glu Ile Asp Ser Ser Asn Leu Ser Leu Trp Ile
                920                 925                 930
Glu Tyr Gly Thr Met Ser Tyr Ala Leu His Ser Phe Ala Ser Arg
                935                 940                 945
Gln Leu Lys Gln Trp Arg Gly Glu Leu Pro Pro Glu Leu Val Gln
                950                 955                 960
Gln Met Glu Gly Arg Arg Asp Ser Met Leu Glu Thr Ala Lys His
                965                 970                 975
Cys Phe Thr Ser Ala Ala Arg Cys Glu Gly Asp Gly Asp Glu Glu
                980                 985                 990
Glu Trp Leu Ile His Tyr Met Leu Gly Lys Val Ala Glu Lys Gln
                995                 1000                1005
Gln Gln Pro Pro Thr Val Tyr Leu Leu His Tyr Arg Gln Ala Gly
               1010                1015                1020
His Tyr Leu His Glu Glu Ala Ala Arg Tyr Pro Lys Lys Ile His
               1025                1030                1035
Tyr His Asn Pro Pro Glu Leu Ala Met Glu Ala Leu Glu Val Tyr
               1040                1045                1050
Phe Arg Leu His Ala Ser Ile Leu Lys Leu Leu Gly Lys Pro Asp
               1055                1060                1065
Ser Gly Val Gly Ala Glu Val Leu Val Asn Phe Met Lys Glu Ala
               1070                1075                1080
Ala Glu Gly Pro Phe Ala Arg Gly Glu Glu Lys Asn Thr Pro Lys
               1085                1090                1095
Ala Ser Glu Lys Glu Lys Ala Cys Leu Val Asp Glu Asp Ser His
               1100                1105                1110
Ser Ser Ala Gly Thr Leu Pro Gly Pro Gly Ala Ser Leu Pro Ser
               1115                1120                1125
Ser Ser Gly Pro Gly Leu Thr Ser Pro Pro Tyr Thr Ala Thr Pro
               1130                1135                1140
Ile Asp His Asp Tyr Val Lys Cys Lys Lys Pro His Gln Gln Ala
               1145                1150                1155
Thr Pro Asp Asp Arg Ser Gln Asp Ser Thr Ala Val Ala Leu Ser
               1160                1165                1170
Asp Ser Ser Ser Thr Gln Asp Phe Phe Asn Glu Pro Thr Ser Leu
               1175                1180                1185
Leu Glu Gly Ser Arg Lys Ser Tyr Thr Glu Lys Arg Leu Pro Ile
               1190                1195                1200
Leu Ser Ser Gln Ala Gly Ala Thr Gly Lys Asp Leu Gln Gly Ala
               1205                1210                1215
Thr Glu Glu Arg Gly Lys Asn Glu Glu Ser Leu Glu Ser Thr Glu
               1220                1225                1230
Gly Phe Arg Ala Ala Glu Gln Gly Val Gln Lys Pro Ala Ala Glu
               1235                1240                1245
Thr Pro Ala Ser Ala Cys Ile Pro Gly Lys Pro Ser Ala Ser Thr
```

```
                1250                1255                1260
Pro Thr Leu Trp Asp Gly Lys Lys Arg Gly Asp Leu Pro Gly Glu
                1265                1270                1275
Pro Val Ala Phe Pro Gln Gly Leu Pro Ala Gly Ala Glu Glu Gln
                1280                1285                1290
Arg Gln Phe Leu Thr Glu Gln Cys Ile Ala Ser Phe Arg Leu Cys
                1295                1300                1305
Leu Ser Arg Phe Pro Gln His Tyr Lys Ser Leu Tyr Arg Leu Ala
                1310                1315                1320
Phe Leu Tyr Thr Tyr Ser Lys Thr His Arg Asn Leu Gln Trp Ala
                1325                1330                1335
Arg Asp Val Leu Leu Gly Ser Ser Ile Pro Trp Gln Gln Leu Gln
                1340                1345                1350
His Met Pro Ala Gln Gly Leu Phe Cys Glu Arg Asn Lys Thr Asn
                1355                1360                1365
Phe Phe Asn Gly Ile Trp Arg Ile Pro Val Asp Glu Ile Asp Arg
                1370                1375                1380
Pro Gly Ser Phe Ala Trp His Met Asn Arg Ser Ile Val Leu Leu
                1385                1390                1395
Leu Lys Val Leu Ala Gln Leu Arg Asp His Ser Thr Leu Leu Lys
                1400                1405                1410
Val Ser Ser Met Leu Gln Arg Thr Pro Asp Gln Gly Lys Lys Tyr
                1415                1420                1425
Leu Arg Asp Ala Asp Arg Gln Val Leu Ala Gln Arg Ala Phe Ile
                1430                1435                1440
Leu Thr Val Lys Val Leu Glu Asp Thr Leu Ser Glu Leu Ala Glu
                1445                1450                1455
Gly Ser Glu Arg Pro Gly Pro Lys Val Cys Gly Leu Pro Gly Ala
                1460                1465                1470
Arg Met Thr Thr Asp Val Ser His Lys Ala Ser Pro Glu Asp Gly
                1475                1480                1485
Gln Glu Gly Leu Pro Gln Pro Lys Lys Pro Pro Leu Ala Asp Gly
                1490                1495                1500
Ser Gly Pro Gly Pro Glu Pro Gly Gly Lys Val Gly Leu Leu Asn
                1505                1510                1515
His Arg Pro Val Ala Met Asp Ala Gly Asp Ser Ala Asp Gln Ser
                1520                1525                1530
Gly Glu Arg Lys Asp Lys Glu Ser Pro Arg Ala Gly Pro Thr Glu
                1535                1540                1545
Pro Met Asp Thr Ser Glu Ala Thr Val Cys His Ser Asp Leu Glu
                1550                1555                1560
Arg Thr Pro Pro Leu Leu Pro Gly Arg Pro Ala Arg Asp Arg Gly
                1565                1570                1575
Pro Glu Ser Arg Pro Thr Glu Leu Ser Leu Glu Glu Leu Ser Ile
                1580                1585                1590
Ser Ala Arg Gln Gln Pro Thr Pro Leu Thr Pro Ala Gln Pro Ala
                1595                1600                1605
Pro Ala Pro Ala Pro Ala Thr Thr Thr Gly Thr Arg Ala Gly Gly
                1610                1615                1620
His Pro Glu Glu Pro Leu Ser Arg Leu Ser Arg Lys Arg Lys Leu
                1625                1630                1635
Leu Glu Asp Thr Glu Ser Gly Lys Thr Leu Leu Leu Asp Ala Tyr
                1640                1645                1650
Arg Val Trp Gln Gln Gly Gln Lys Gly Val Ala Tyr Asp Leu Gly
                1655                1660                1665
Arg Val Glu Arg Ile Met Ser Glu Thr Tyr Met Leu Ile Lys Gln
```

```
               1670                1675                1680
Val Asp Glu Glu Ala Ala Leu Glu Gln Ala Val Lys Phe Cys Gln
               1685                1690                1695
Val His Leu Gly Ala Ala Ala Gln Arg Gln Ala Ser Gly Asp Thr
               1700                1705                1710
Pro Thr Thr Pro Lys His Pro Lys Asp Ser Arg Glu Asn Phe Phe
               1715                1720                1725
Pro Val Thr Val Val Pro Thr Ala Pro Asp Pro Val Pro Ala Asp
               1730                1735                1740
Ser Val Gln Arg Pro Ser Asp Ala His Thr Lys Pro Arg Pro Ala
               1745                1750                1755
Leu Ala Ala Ala Thr Thr Ile Ile Thr Cys Pro Pro Ser Ala Ser
               1760                1765                1770
Ala Ser Thr Leu Asp Gln Ser Lys Asp Pro Gly Pro Pro Arg Pro
               1775                1780                1785
His Arg Pro Glu Ala Thr Pro Ser Met Ala Ser Leu Gly Pro Glu
               1790                1795                1800
Gly Glu Glu Leu Ala Arg Val Ala Glu Gly Thr Ser Phe Pro Pro
               1805                1810                1815
Gln Glu Pro Arg His Ser Pro Gln Val Lys Met Ala Pro Thr Ser
               1820                1825                1830
Ser Pro Ala Glu Pro His Cys Trp Pro Ala Glu Ala Ala Leu Gly
               1835                1840                1845
Thr Gly Ala Glu Pro Thr Cys Ser Gln Ala Ala Ser Ser Lys Ala
               1850                1855                1860
Pro Ser Ser Gly Ser Ala Gln Pro Pro Glu Gly His Pro Gly Lys
               1865                1870                1875
Pro Glu Pro Ser Arg Ala Lys Ser Arg Pro Leu Pro Asn Met Pro
               1880                1885                1890
Lys Leu Val Ile Pro Ser Ala Ala Thr Lys Phe Pro Pro Glu Ile
               1895                1900                1905
Thr Val Thr Pro Pro Thr Pro Thr Leu Leu Ser Pro Lys Gly Ser
               1910                1915                1920
Ile Ser Glu Glu Thr Lys Gln Lys Leu Lys Ser Ala Ile Leu Ser
               1925                1930                1935
Ala Gln Ser Ala Ala Asn Val Arg Lys Glu Ser Leu Cys Gln Pro
               1940                1945                1950
Ala Leu Glu Val Leu Glu Thr Ser Ser Gln Glu Ser Ser Leu Glu
               1955                1960                1965
Ser Glu Thr Asp Glu Asp Asp Asp Tyr Met Asp Ile
               1970                1975

<210> 4
<211> 279
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505083CD1

<400> 4
Met Pro Val Pro Ala Ser Trp Pro His Leu Pro Ser Pro Phe Leu
 1               5                  10                  15
Leu Met Thr Leu Leu Leu Gly Arg Leu Thr Ala Lys Pro Ser Ala
                20                  25                  30
```

```
Pro Val Val Ser Gly Pro Ala Val Arg Ala Thr Pro Glu His Thr
                35                  40                  45
Val Ser Phe Thr Cys Glu Ser His Gly Phe Ser Pro Arg Asp Ile
                50                  55                  60
Thr Leu Lys Trp Phe Lys Asn Gly Asn Glu Leu Ser Asp Phe Gln
                65                  70                  75
Thr Asn Val Asp Pro Ala Gly Asp Ser Val Ser Tyr Ser Ile His
                80                  85                  90
Ser Thr Ala Arg Val Val Leu Thr Arg Gly Asp Val His Ser Gln
                95                  100                 105
Val Ile Cys Glu Met Ala His Ile Thr Leu Gln Gly Asp Pro Leu
                110                 115                 120
Arg Gly Thr Ala Asn Leu Ser Glu Ala Ile Arg Val Pro Pro Thr
                125                 130                 135
Leu Glu Val Thr Gln Gln Pro Met Arg Ala Glu Asn Gln Ala Asn
                140                 145                 150
Val Thr Cys Gln Val Ser Asn Phe Tyr Pro Arg Gly Leu Gln Leu
                155                 160                 165
Thr Trp Leu Glu Asn Gly Asn Val Ser Arg Thr Glu Thr Ala Ser
                170                 175                 180
Thr Leu Ile Glu Asn Lys Asp Gly Thr Tyr Asn Trp Met Ser Trp
                185                 190                 195
Leu Leu Val Asn Thr Cys Ala His Arg Asp Asp Val Val Leu Thr
                200                 205                 210
Cys Gln Val Glu His Asp Gly Gln Gln Ala Val Ser Lys Ser Tyr
                215                 220                 225
Ala Leu Glu Ile Ser Ala His Gln Lys Glu His Gly Ser Asp Ile
                230                 235                 240
Thr His Glu Pro Ala Leu Ala Pro Thr Ala Pro Leu Leu Val Ala
                245                 250                 255
Leu Leu Leu Gly Pro Lys Leu Leu Leu Val Val Gly Val Ser Ala
                260                 265                 270
Ile Tyr Ile Cys Trp Lys Gln Lys Ala
                275

<210> 5
<211> 1207
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505866CD1

<400> 5
Met Arg Ser Arg Ser Asn Ser Gly Val Arg Leu Asp Gly Tyr Ala
  1               5                   10                  15
Arg Leu Val Gln Gln Thr Ile Leu Cys Tyr Gln Asn Pro Val Thr
                20                  25                  30
Gly Leu Leu Ser Ala Ser His Glu Gln Lys Asp Ala Trp Val Arg
                35                  40                  45
Asp Asn Ile Tyr Ser Ile Leu Ala Val Trp Gly Leu Gly Met Ala
                50                  55                  60
Tyr Arg Lys Asn Ala Asp Arg Asp Glu Asp Lys Ala Lys Ala Tyr
                65                  70                  75
Glu Leu Glu Gln Asn Val Val Lys Leu Met Arg Gly Leu Leu Gln
```

```
                80                  85                  90
Cys Met Met Arg Gln Val Ala Lys Val Glu Lys Phe Lys His Thr
                95                 100                 105
Gln Ser Thr Lys Asp Ser Leu His Ala Lys Tyr Asn Thr Ala Thr
                110                 115                 120
Cys Gly Thr Val Val Gly Asp Asp Gln Trp Gly His Leu Gln Val
                125                 130                 135
Asp Ala Thr Ser Leu Phe Leu Leu Phe Leu Ala Gln Met Thr Ala
                140                 145                 150
Ser Gly Leu Arg Ile Ile Phe Thr Leu Asp Glu Val Ala Phe Ile
                155                 160                 165
Gln Asn Leu Val Phe Tyr Ile Glu Ala Ala Tyr Lys Val Ala Asp
                170                 175                 180
Tyr Gly Met Trp Glu Arg Gly Asp Lys Thr Asn Gln Gly Ile Pro
                185                 190                 195
Glu Leu Asn Ala Ser Ser Val Gly Met Ala Lys Ala Ala Leu Glu
                200                 205                 210
Ala Ile Asp Glu Leu Asp Leu Phe Gly Ala His Gly Gly Arg Lys
                215                 220                 225
Ser Val Ile His Val Leu Pro Asp Glu Val Glu His Cys Gln Ser
                230                 235                 240
Ile Leu Phe Ser Met Leu Pro Arg Ala Ser Thr Ser Lys Glu Ile
                245                 250                 255
Asp Ala Gly Leu Leu Ser Ile Ile Ser Phe Pro Ala Phe Ala Val
                260                 265                 270
Glu Asp Val Asn Leu Val Asn Val Thr Lys Asn Glu Ile Ile Ser
                275                 280                 285
Lys Leu Gln Gly Arg Tyr Gly Cys Cys Arg Phe Leu Arg Asp Gly
                290                 295                 300
Tyr Lys Thr Pro Arg Glu Asp Pro Asn Arg Leu His Tyr Asp Pro
                305                 310                 315
Ala Glu Leu Lys Leu Phe Glu Asn Ile Glu Cys Glu Trp Pro Val
                320                 325                 330
Phe Trp Thr Tyr Phe Ile Ile Asp Gly Val Phe Ser Gly Asp Ala
                335                 340                 345
Val Gln Val Gln Glu Tyr Arg Glu Ala Leu Glu Gly Ile Leu Ile
                350                 355                 360
Arg Gly Lys Asn Gly Ile Arg Leu Val Pro Glu Leu Tyr Ala Val
                365                 370                 375
Pro Pro Asn Lys Val Asp Glu Glu Tyr Lys Asn Pro His Thr Val
                380                 385                 390
Asp Arg Val Pro Met Gly Lys Val Pro His Leu Trp Gly Gln Ser
                395                 400                 405
Leu Tyr Ile Leu Ser Ser Leu Leu Ala Glu Gly Phe Leu Ala Ala
                410                 415                 420
Gly Glu Ile Asp Pro Leu Asn Arg Arg Phe Ser Thr Ser Val Lys
                425                 430                 435
Pro Asp Val Val Val Gln Val Thr Val Leu Ala Glu Asn Asn His
                440                 445                 450
Ile Lys Asp Leu Leu Arg Lys His Gly Val Asn Val Gln Ser Ile
                455                 460                 465
Ala Asp Ile His Pro Ile Gln Val Gln Pro Gly Arg Ile Leu Ser
                470                 475                 480
His Ile Tyr Ala Lys Leu Gly Arg Asn Lys Asn Met Asn Leu Ser
                485                 490                 495
Gly Arg Pro Tyr Arg His Ile Gly Val Leu Gly Thr Ser Lys Leu
```

```
                500                 505                 510
Tyr Val Ile Arg Asn Gln Ile Phe Thr Phe Thr Pro Gln Phe Thr
                515                 520                 525
Asp Gln His His Phe Tyr Leu Ala Leu Asp Asn Glu Met Ile Val
                530                 535                 540
Glu Met Leu Arg Ile Glu Leu Ala Tyr Leu Cys Thr Cys Trp Arg
                545                 550                 555
Met Thr Gly Arg Pro Thr Leu Thr Phe Pro Ile Ser Arg Thr Met
                560                 565                 570
Leu Thr Asn Asp Gly Ser Asp Ile His Ser Ala Val Leu Ser Thr
                575                 580                 585
Ile Arg Lys Leu Glu Asp Gly Tyr Phe Gly Gly Ala Arg Val Lys
                590                 595                 600
Leu Gly Asn Leu Ser Glu Phe Leu Thr Thr Ser Phe Tyr Thr Tyr
                605                 610                 615
Leu Thr Phe Leu Asp Pro Asp Cys Asp Glu Lys Leu Phe Asp Asn
                620                 625                 630
Ala Ser Glu Gly Thr Phe Ser Pro Asp Ser Asp Ser Asp Leu Val
                635                 640                 645
Gly Tyr Leu Glu Asp Thr Cys Asn Gln Glu Ser Gln Asp Glu Leu
                650                 655                 660
Asp His Tyr Ile Asn His Leu Leu Gln Ser Thr Ser Leu Arg Ser
                665                 670                 675
Tyr Leu Pro Pro Leu Cys Lys Asn Thr Glu Asp Arg His Val Phe
                680                 685                 690
Ser Ala Ile His Ser Thr Arg Asp Ile Leu Ser Val Met Ala Lys
                695                 700                 705
Ala Lys Gly Leu Glu Val Pro Phe Val Pro Met Thr Leu Pro Thr
                710                 715                 720
Lys Val Leu Ser Ala His Arg Lys Ser Leu Asn Leu Val Asp Ser
                725                 730                 735
Pro Gln Pro Leu Leu Glu Lys Val Pro Glu Ser Asp Phe Gln Trp
                740                 745                 750
Pro Arg Asp Asp His Gly Asp Val Asp Cys Glu Lys Leu Val Glu
                755                 760                 765
Gln Leu Lys Asp Cys Ser Asn Leu Gln Asp Gln Ala Asp Ile Leu
                770                 775                 780
Tyr Ile Leu Tyr Val Ile Lys Gly Pro Ser Trp Asp Thr Asn Leu
                785                 790                 795
Ser Gly Gln His Gly Val Thr Val Gln Asn Leu Leu Gly Glu Leu
                800                 805                 810
Tyr Gly Lys Ala Gly Leu Asn Gln Glu Trp Gly Leu Ile Arg Tyr
                815                 820                 825
Ile Ser Gly Leu Leu Arg Lys Lys Val Glu Val Leu Ala Glu Ala
                830                 835                 840
Cys Thr Asp Leu Leu Ser His Gln Lys Gln Leu Thr Val Gly Leu
                845                 850                 855
Pro Pro Glu Pro Arg Glu Lys Ile Ile Ser Ala Pro Leu Pro Pro
                860                 865                 870
Glu Glu Leu Thr Lys Leu Ile Tyr Glu Ala Ser Gly Gln Asp Ile
                875                 880                 885
Ser Ile Ala Val Leu Thr Gln Glu Ile Val Val Tyr Leu Ala Met
                890                 895                 900
Tyr Val Arg Ala Gln Pro Ser Leu Phe Val Glu Met Leu Arg Leu
                905                 910                 915
Arg Ile Gly Leu Ile Ile Gln Val Met Ala Thr Glu Leu Ala Arg
```

```
                920                     925                     930
Ser Leu Asn Cys Ser Gly Glu Glu Ala Ser Glu Ser Leu Met Asn
                935                     940                     945
Leu Ser Pro Phe Asp Met Lys Asn Leu Leu His His Ile Leu Ser
                950                     955                     960
Gly Lys Glu Phe Gly Val Glu Arg Ser Val Arg Pro Ile His Ser
                965                     970                     975
Ser Thr Ser Ser Pro Thr Ile Ser Ile His Glu Val Gly His Thr
                980                     985                     990
Gly Val Thr Lys Thr Glu Arg Ser Gly Ile Asn Arg Leu Arg Ser
                995                     1000                    1005
Glu Met Lys Gln Arg Ser Ser Thr Pro Ser Ser Pro Thr Gly Thr
                1010                    1015                    1020
Ser Ser Ser Asp Ser Gly Gly His His Ile Gly Trp Gly Glu Arg
                1025                    1030                    1035
Gln Gly Gln Trp Leu Arg Arg Arg Arg Leu Asp Gly Ala Ile Asn
                1040                    1045                    1050
Arg Val Pro Val Gly Phe Tyr Gln Arg Val Trp Lys Ile Leu Gln
                1055                    1060                    1065
Lys Cys His Gly Leu Ser Ile Asp Gly Tyr Val Leu Pro Ser Ser
                1070                    1075                    1080
Thr Thr Arg Glu Met Thr Pro His Glu Ile Lys Phe Ala Val His
                1085                    1090                    1095
Val Glu Ser Val Leu Asn Arg Val Pro Gln Pro Glu Tyr Arg Gln
                1100                    1105                    1110
Leu Leu Val Glu Ala Ile Met Val Leu Thr Leu Leu Ser Asp Thr
                1115                    1120                    1125
Glu Met Thr Ser Ile Gly Gly Ile Ile His Val Asp Gln Ile Val
                1130                    1135                    1140
Gln Met Ala Ser Gln Leu Phe Leu Gln Asp Gln Val Ser Ile Gly
                1145                    1150                    1155
Ala Met Asp Thr Leu Glu Lys Asp Gln Ala Thr Gly Ile Cys His
                1160                    1165                    1170
Phe Phe Tyr Asp Ser Ala Pro Ser Gly Ala Tyr Gly Thr Met Thr
                1175                    1180                    1185
Tyr Leu Thr Arg Ala Val Ala Ser Tyr Leu Gln Glu Leu Leu Pro
                1190                    1195                    1200
Asn Ser Gly Cys Gln Met Gln
                1205

<210> 6
<211> 1328
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503214CD1

<400> 6
Met Lys Trp Val Gly Asp Thr Gly Val Gly Gly Asn Ile Pro Pro
  1               5                  10                  15
Ser Phe Thr Thr Pro Gly Leu Ser Ser Arg Pro Gly Ala Met Val
                 20                  25                  30
Ala Asp Arg Ser Arg Trp Pro Leu Ala Gln Gly Lys Gly Ala Gln
                 35                  40                  45
```

```
Ala Gly Thr Trp Arg Ala Ala Val Glu Cys Ser Gly Arg Gly Leu
                50                  55                  60
Gly Ala Ala Ser Glu Ser Pro Gln Cys Pro Pro Pro Pro Gly Val
                65                  70                  75
Glu Gly Ala Ala Gly Pro Ala Glu Pro Asp Gly Ala Ala Glu Gly
                80                  85                  90
Ala Ala Gly Gly Ser Gly Glu Gly Glu Ser Gly Gly Gly Pro Arg
                95                  100                 105
Arg Ala Leu Arg Ala Val Tyr Val Arg Ser Glu Ser Ser Gln Gly
                110                 115                 120
Gly Ala Ala Gly Gly Pro Glu Ala Gly Ala Arg Gln Cys Leu Leu
                125                 130                 135
Arg Ala Cys Glu Ala Glu Gly Ala His Leu Thr Ser Val Pro Phe
                140                 145                 150
Gly Glu Leu Asp Phe Gly Glu Thr Ala Val Leu Asp Ala Phe Tyr
                155                 160                 165
Asp Ala Asp Val Ala Val Val Asp Met Ser Asp Val Ser Arg Gln
                170                 175                 180
Pro Ser Leu Phe Tyr His Leu Gly Val Arg Glu Ser Phe Asp Met
                185                 190                 195
Ala Asn Asn Val Ile Leu Tyr His Asp Thr Asp Ala Asp Thr Ala
                200                 205                 210
Leu Ser Leu Lys Asp Met Val Thr Gln Lys Asn Thr Ala Ser Ser
                215                 220                 225
Gly Asn Tyr Tyr Phe Ile Pro Tyr Ile Val Thr Pro Cys Thr Asp
                230                 235                 240
Tyr Phe Cys Cys Glu Ser Asp Ala Gln Arg Arg Ala Ser Glu Tyr
                245                 250                 255
Met Gln Pro Asn Trp Asp Asn Ile Leu Gly Pro Leu Cys Met Pro
                260                 265                 270
Leu Val Asp Arg Phe Ile Ser Leu Leu Lys Asp Ile His Val Thr
                275                 280                 285
Ser Cys Val Tyr Tyr Lys Glu Thr Leu Leu Asn Asp Ile Arg Lys
                290                 295                 300
Ala Arg Glu Lys Tyr Gln Gly Glu Glu Leu Ala Lys Glu Leu Ala
                305                 310                 315
Arg Ile Lys Leu Arg Met Asp Asn Thr Glu Val Leu Thr Ser Asp
                320                 325                 330
Ile Ile Ile Asn Leu Leu Leu Ser Tyr Arg Asp Ile Gln Asp Tyr
                335                 340                 345
Asp Ala Met Val Lys Leu Val Glu Thr Leu Glu Met Leu Pro Thr
                350                 355                 360
Cys Asp Leu Ala Asp Gln His Asn Ile Lys Phe His Tyr Ala Phe
                365                 370                 375
Ala Leu Asn Arg Arg Asn Ser Thr Gly Asp Arg Glu Lys Ala Leu
                380                 385                 390
Gln Ile Met Leu Gln Val Leu Gln Ser Cys Asp His Pro Gly Pro
                395                 400                 405
Asp Met Phe Cys Leu Cys Gly Arg Ile Tyr Lys Asp Ile Phe Leu
                410                 415                 420
Asp Ser Asp Cys Lys Asp Asp Thr Ser Arg Asp Ser Ala Ile Glu
                425                 430                 435
Trp Tyr Arg Lys Gly Phe Glu Leu Gln Ser Ser Leu Tyr Ser Gly
                440                 445                 450
Ile Asn Leu Ala Val Leu Leu Ile Val Ala Gly Gln Gln Phe Glu
                455                 460                 465
```

```
Thr Ser Leu Glu Leu Arg Lys Ile Gly Val Arg Leu Asn Ser Leu
                470                 475                 480
Leu Gly Arg Lys Gly Ser Leu Glu Lys Met Asn Asn Tyr Trp Asp
                485                 490                 495
Val Gly Gln Phe Phe Ser Val Ser Met Leu Ala His Asp Val Gly
                500                 505                 510
Lys Ala Val Gln Ala Ala Glu Arg Leu Phe Lys Leu Lys Pro Pro
                515                 520                 525
Val Trp Tyr Leu Arg Ser Leu Val Gln Asn Leu Leu Leu Ile Arg
                530                 535                 540
Arg Phe Lys Lys Thr Ile Ile Glu His Ser Pro Arg Gln Glu Arg
                545                 550                 555
Leu Asn Phe Trp Leu Asp Ile Ile Phe Glu Ala Thr Asn Glu Val
                560                 565                 570
Thr Asn Gly Leu Arg Phe Pro Val Leu Val Ile Glu Pro Thr Lys
                575                 580                 585
Val Tyr Gln Pro Ser Tyr Val Ser Ile Asn Asn Glu Ala Glu Glu
                590                 595                 600
Arg Thr Val Ser Leu Trp His Val Ser Pro Thr Glu Met Lys Gln
                605                 610                 615
Met His Glu Trp Asn Phe Thr Ala Ser Ser Ile Lys Gly Ile Arg
                620                 625                 630
Phe Phe Ser Leu Val Lys Glu Met Ile Thr Asn Thr Ala Gly Ser
                635                 640                 645
Thr Val Glu Leu Glu Gly Glu Thr Asp Gly Asp Thr Leu Glu Tyr
                650                 655                 660
Glu Tyr Asp His Asp Ala Asn Gly Glu Arg Val Val Leu Gly Lys
                665                 670                 675
Gly Thr Tyr Gly Ile Val Tyr Ala Gly Arg Asp Leu Ser Asn Gln
                680                 685                 690
Val Arg Ile Ala Ile Lys Glu Ile Pro Glu Arg Asp Ser Arg Tyr
                695                 700                 705
Ser Gln Pro Leu His Glu Glu Ile Ala Leu His Lys Tyr Leu Lys
                710                 715                 720
His Arg Asn Ile Val Gln Tyr Leu Gly Ser Val Ser Glu Asn Gly
                725                 730                 735
Tyr Ile Lys Ile Phe Met Glu Gln Val Pro Gly Gly Ser Leu Ser
                740                 745                 750
Ala Leu Leu Arg Ser Lys Trp Gly Pro Met Lys Glu Pro Thr Ile
                755                 760                 765
Lys Phe Tyr Thr Lys Gln Ile Leu Glu Gly Leu Lys Tyr Leu His
                770                 775                 780
Glu Asn Gln Ile Val His Arg Asp Ile Lys Gly Asp Asn Val Leu
                785                 790                 795
Val Asn Thr Tyr Ser Gly Val Val Lys Ile Ser Asp Phe Gly Thr
                800                 805                 810
Ser Lys Arg Leu Ala Gly Val Asn Pro Cys Thr Glu Thr Phe Thr
                815                 820                 825
Gly Thr Leu Gln Tyr Met Ala Pro Glu Ile Ile Asp Gln Gly Pro
                830                 835                 840
Arg Gly Tyr Gly Ala Pro Ala Asp Ile Trp Ser Leu Gly Cys Thr
                845                 850                 855
Ile Ile Glu Met Ala Thr Ser Lys Pro Pro Phe His Glu Leu Gly
                860                 865                 870
Glu Pro Gln Ala Ala Met Phe Lys Val Gly Met Phe Lys Ile His
                875                 880                 885
```

```
Pro Glu Ile Pro Glu Ala Leu Ser Ala Glu Ala Arg Ala Phe Ile
                890                 895                 900
Leu Ser Cys Phe Glu Pro Asp Pro His Lys Arg Ala Thr Thr Ala
                905                 910                 915
Glu Leu Leu Arg Glu Gly Phe Leu Arg Gln Val Asn Lys Gly Lys
                920                 925                 930
Lys Asn Arg Ile Ala Phe Lys Pro Ser Glu Gly Pro Arg Gly Val
                935                 940                 945
Val Leu Ala Leu Pro Thr Gln Gly Glu Pro Met Ala Thr Ser Ser
                950                 955                 960
Ser Glu His Gly Ser Val Ser Pro Asp Ser Asp Ala Gln Pro Asp
                965                 970                 975
Ala Leu Phe Glu Arg Thr Arg Ala Pro Arg His His Leu Gly His
                980                 985                 990
Leu Leu Ser Val Pro Asp Glu Ser Ser Ala Leu Glu Asp Arg Gly
                995                1000                1005
Leu Ala Ser Ser Pro Glu Asp Arg Asp Gln Gly Leu Phe Leu Leu
               1010                1015                1020
Arg Lys Asp Ser Glu Arg Arg Ala Ile Leu Tyr Lys Ile Leu Trp
               1025                1030                1035
Glu Glu Gln Asn Gln Val Ala Ser Asn Leu Gln Glu Cys Val Ala
               1040                1045                1050
Gln Ser Ser Glu Glu Leu His Leu Ser Val Gly His Ile Lys Gln
               1055                1060                1065
Ile Ile Gly Ile Leu Arg Asp Phe Ile Arg Ser Pro Glu His Arg
               1070                1075                1080
Val Met Ala Thr Thr Ile Ser Lys Leu Lys Val Asp Leu Asp Phe
               1085                1090                1095
Asp Ser Ser Ser Ile Ser Gln Ile His Leu Val Leu Phe Gly Phe
               1100                1105                1110
Gln Asp Ala Val Asn Lys Ile Leu Arg Asn His Leu Ile Arg Pro
               1115                1120                1125
His Trp Met Phe Ala Met Asp Asn Ile Ile Arg Arg Ala Val Gln
               1130                1135                1140
Ala Ala Val Thr Ile Leu Ile Pro Glu Leu Arg Ala His Phe Glu
               1145                1150                1155
Pro Thr Cys Glu Thr Glu Gly Val Asp Lys Asp Met Asp Glu Ala
               1160                1165                1170
Glu Glu Gly Tyr Pro Pro Ala Thr Gly Pro Gly Gln Glu Ala Gln
               1175                1180                1185
Pro His Gln Gln His Leu Ser Leu Gln Leu Gly Glu Leu Arg Gln
               1190                1195                1200
Glu Thr Asn Arg Leu Leu Glu His Leu Val Glu Lys Glu Arg Glu
               1205                1210                1215
Tyr Gln Asn Leu Leu Arg Gln Thr Leu Glu Gln Lys Thr Gln Glu
               1220                1225                1230
Leu Tyr His Leu Gln Leu Lys Leu Lys Ser Asn Cys Ile Thr Glu
               1235                1240                1245
Asn Pro Ala Gly Pro Tyr Gly Gln Arg Thr Asp Lys Glu Leu Ile
               1250                1255                1260
Asp Trp Leu Arg Leu Gln Gly Ala Asp Ala Lys Thr Ile Glu Lys
               1265                1270                1275
Ile Val Glu Glu Gly Tyr Thr Leu Ser Asp Ile Leu Asn Glu Ile
               1280                1285                1290
Thr Lys Glu Asp Leu Arg Tyr Leu Arg Leu Arg Gly Gly Leu Leu
               1295                1300                1305
```

Cys Arg Leu Trp Ser Ala Val Ser Gln Tyr Arg Arg Ala Gln Glu
1310 1315 1320

Ala Ser Glu Thr Lys Asp Lys Ala
1325

<210> 7
<211> 603
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7495312CD1

<400> 7
Met Gly Asn Tyr Lys Ser Arg Pro Thr Gln Thr Cys Thr Asp Glu
1 5 10 15
Trp Lys Lys Lys Val Ser Glu Ser Tyr Val Ile Thr Ile Glu Arg
20 25 30
Leu Glu Asp Asp Leu Gln Ile Lys Glu Lys Glu Leu Thr Glu Leu
35 40 45
Arg Asn Ile Phe Gly Ser Asp Glu Ala Phe Ser Lys Val Asn Leu
50 55 60
Asn Tyr Arg Thr Glu Asn Gly Leu Ser Leu Leu His Leu Cys Cys
65 70 75
Ile Cys Gly Gly Lys Lys Ser His Ile Arg Thr Leu Met Leu Lys
80 85 90
Gly Leu Arg Pro Ser Arg Leu Thr Arg Asn Gly Phe Thr Ala Leu
95 100 105
His Leu Ala Val Tyr Lys Asp Asn Ala Glu Leu Ile Thr Ser Leu
110 115 120
Leu His Ser Gly Ala Asp Ile Gln Gln Val Gly Tyr Gly Gly Leu
125 130 135
Thr Ala Leu His Ile Ala Thr Ile Ala Gly His Leu Glu Ala Ala
140 145 150
Asp Val Leu Leu Gln His Gly Ala Asn Val Asn Ile Gln Asp Ala
155 160 165
Val Phe Phe Thr Pro Leu His Ile Ala Ala Tyr Tyr Gly His Glu
170 175 180
Gln Val Thr Arg Leu Leu Leu Lys Phe Gly Ala Asp Val Asn Val
185 190 195
Ser Gly Glu Val Gly Asp Arg Pro Leu His Leu Ala Ser Ala Lys
200 205 210
Gly Phe Leu Asn Ile Ala Lys Leu Leu Met Glu Glu Gly Ser Lys
215 220 225
Ala Asp Val Asn Ala Gln Asp Asn Glu Asp His Val Pro Leu His
230 235 240
Phe Cys Ser Arg Phe Gly His His Asp Ile Val Lys Tyr Leu Leu
245 250 255
Gln Ser Asp Leu Glu Val Gln Pro His Val Val Asn Ile Tyr Gly
260 265 270
Asp Thr Pro Leu His Leu Ala Cys Tyr Asn Gly Lys Phe Glu Val
275 280 285
Ala Lys Glu Ile Ile Gln Ile Ser Gly Thr Glu Ser Leu Thr Lys
290 295 300
Glu Asn Ile Phe Ser Glu Thr Ala Phe His Ser Ala Cys Thr Tyr

305 310 315
Gly Lys Ser Ile Asp Leu Val Lys Phe Leu Leu Asp Gln Asn Val
320 325 330
Ile Asn Ile Asn His Gln Gly Arg Asp Gly His Thr Gly Leu His
335 340 345
Ser Ala Cys Tyr His Gly His Ile Arg Leu Val Gln Phe Leu Leu
350 355 360
Asp Asn Gly Ala Asp Met Asn Leu Val Ala Cys Asp Pro Ser Arg
365 370 375
Ser Ser Gly Glu Lys Asp Glu Gln Thr Cys Leu Met Trp Ala Tyr
380 385 390
Glu Lys Gly His Asp Ala Ile Val Thr Leu Leu Lys His Tyr Lys
395 400 405
Arg Pro Gln Asp Glu Leu Pro Cys Asn Glu Tyr Ser Gln Pro Gly
410 415 420
Gly Asp Gly Ser Tyr Val Ser Val Pro Ser Pro Leu Gly Lys Ile
425 430 435
Lys Ser Met Thr Lys Glu Lys Ala Asp Ile Leu Leu Leu Arg Ala
440 445 450
Gly Leu Pro Ser His Phe His Leu Gln Leu Ser Glu Ile Glu Phe
455 460 465
His Glu Ile Ile Gly Ser Gly Ser Phe Gly Lys Val Tyr Lys Gly
470 475 480
Arg Cys Arg Asn Lys Ile Val Ala Ile Lys Arg Tyr Arg Ala Asn
485 490 495
Thr Tyr Cys Ser Lys Ser Asp Val Asp Met Phe Cys Arg Glu Val
500 505 510
Ser Ile Leu Cys Gln Leu Asn His Pro Cys Val Ile Gln Phe Val
515 520 525
Gly Ala Cys Leu Asn Asp Pro Ser Gln Phe Ala Ile Val Thr Gln
530 535 540
Tyr Ile Ser Gly Gly Ser Leu Phe Ser Leu Leu His Glu Gln Lys
545 550 555
Arg Ile Leu Asp Leu Gln Ser Lys Leu Ile Ile Ala Val Asp Val
560 565 570
Ala Lys Gly Met Glu Tyr Leu His Asn Leu Thr Gln Pro Ile Ile
575 580 585
His Arg Asp Leu Asn Arg Cys Cys Thr Gly Trp Leu Ser Cys Tyr
590 595 600
His Pro Asp

<210> 8
<211> 160
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506732CD1

<400> 8
Met Ser Gly Pro Arg Pro Val Val Leu Ser Gly Pro Ser Gly Ala
1 5 10 15
Gly Lys Ser Thr Leu Leu Lys Arg Leu Leu Gln Glu His Ser Gly
20 25 30

```
Ile Phe Gly Phe Ser Val Ser His Thr Thr Arg Asn Leu Tyr Gly
                35                  40                  45
Thr Ser Lys Val Ala Val Gln Ala Val Gln Ala Met Asn Arg Ile
                50                  55                  60
Cys Val Leu Asp Val Asp Leu Gln Gly Val Arg Asn Ile Lys Ala
                65                  70                  75
Thr Asp Leu Arg Pro Ile Tyr Ile Ser Val Gln Pro Pro Ser Leu
                80                  85                  90
His Val Leu Glu Gln Arg Leu Arg Gln Arg Asn Thr Glu Thr Glu
                95                  100                 105
Glu Ser Leu Val Lys Arg Leu Ala Ala Ala Gln Ala Asp Met Glu
                110                 115                 120
Ser Ser Lys Glu Pro Gly Leu Phe Asp Val Val Ile Ile Asn Asp
                125                 130                 135
Ser Leu Asp Gln Ala Tyr Ala Glu Leu Lys Glu Ala Leu Ser Glu
                140                 145                 150
Glu Ile Lys Lys Ala Gln Arg Thr Gly Gly
                155                 160

<210> 9
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506736CD1

<400> 9
Met Ser Gly Pro Arg Pro Val Val Leu Ser Gly Pro Ser Gly Ala
  1               5                   10                  15
Gly Lys Ser Thr Leu Leu Lys Arg Leu Leu Gln Glu His Ser Gly
                20                  25                  30
Ile Phe Gly Phe Ser Val Ser His Thr Thr Arg Asn Pro Arg Pro
                35                  40                  45
Gly Glu Glu Asn Gly Lys Asp Tyr Tyr Phe Val Thr Arg Glu Val
                50                  55                  60
Met Gln Arg Asp Ile Ala Ala Gly Asp Phe Ile Glu His Ala Glu
                65                  70                  75
Phe Ser Arg Arg Ala Ala Arg Ser Pro Ala Cys Leu Met Trp Ser
                80                  85                  90
Ser Leu Thr Thr Ala Trp Thr Arg Pro Thr Gln Ser
                95                  100

<210> 10
<211> 2191
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507121CD1

<400> 10
Met Ile Arg Ile Ala Ala Leu Asn Ala Ser Ser Thr Ile Glu Asp
  1               5                   10                  15
```

```
Asp His Glu Gly Ser Phe Lys Ser His Lys Thr Gln Thr Lys Glu
                20                  25                  30
Ala Gln Glu Ala Glu Ala Phe Ala Leu Tyr His Lys Ala Leu Asp
                35                  40                  45
Leu Gln Lys His Asp Arg Phe Glu Glu Ser Ala Lys Ala Tyr His
                50                  55                  60
Glu Leu Leu Glu Ala Ser Leu Leu Arg Glu Ala Val Ser Ser Gly
                65                  70                  75
Asp Glu Lys Glu Gly Leu Lys His Pro Gly Leu Ile Leu Lys Tyr
                80                  85                  90
Ser Thr Tyr Lys Asn Leu Ala Gln Leu Ala Ala Gln Arg Glu Asp
                95                  100                 105
Leu Glu Thr Ala Met Glu Phe Tyr Leu Glu Ala Val Met Leu Asp
                110                 115                 120
Ser Thr Asp Val Asn Leu Trp Tyr Lys Ile Gly His Val Ala Leu
                125                 130                 135
Arg Leu Ile Arg Ile Pro Leu Ala Arg His Ala Phe Glu Glu Gly
                140                 145                 150
Leu Arg Cys Asn Pro Asp His Trp Pro Cys Leu Asp Asn Leu Ile
                155                 160                 165
Thr Val Leu Tyr Thr Leu Ser Asp Tyr Thr Thr Cys Leu Tyr Phe
                170                 175                 180
Ile Cys Lys Ala Leu Glu Lys Asp Cys Arg Tyr Ser Lys Gly Leu
                185                 190                 195
Val Leu Lys Glu Lys Ile Phe Glu Glu Gln Pro Cys Leu Arg Lys
                200                 205                 210
Asp Ser Leu Arg Met Phe Leu Lys Cys Asp Met Ser Ile His Asp
                215                 220                 225
Val Ser Val Ser Ala Ala Glu Thr Gln Ala Ile Val Asp Glu Ala
                230                 235                 240
Leu Gly Leu Arg Lys Lys Arg Gln Ala Leu Ile Val Arg Glu Lys
                245                 250                 255
Glu Pro Asp Leu Lys Leu Val Gln Pro Ile Pro Phe Phe Thr Trp
                260                 265                 270
Lys Cys Leu Gly Glu Ser Leu Leu Ala Met Tyr Asn His Leu Thr
                275                 280                 285
Thr Cys Glu Pro Pro Arg Pro Ser Leu Gly Lys Arg Ile Asp Leu
                290                 295                 300
Ser Asp Tyr Gln Asp Pro Ser Gln Pro Leu Glu Ser Ser Met Val
                305                 310                 315
Val Thr Pro Val Asn Val Ile Gln Pro Ser Thr Val Ser Thr Asn
                320                 325                 330
Pro Ala Val Ala Val Ala Glu Pro Val Val Ser Tyr Thr Ser Val
                335                 340                 345
Ala Thr Thr Ser Phe Pro Leu His Ser Pro Gly Leu Leu Glu Thr
                350                 355                 360
Gly Ala Pro Val Gly Asp Ile Ser Gly Gly Asp Lys Ser Lys Lys
                365                 370                 375
Gly Val Lys Arg Lys Lys Ile Ser Glu Glu Ser Gly Glu Thr Ala
                380                 385                 390
Lys Arg Arg Ser Ala Arg Val Arg Asn Thr Lys Cys Lys Lys Glu
                395                 400                 405
Glu Lys Val Asp Phe Gln Glu Leu Leu Met Lys Phe Leu Pro Ser
                410                 415                 420
Arg Leu Arg Lys Leu Asp Pro Glu Glu Glu Asp Asp Ser Phe Asn
                425                 430                 435
```

Asn Tyr Glu Val Gln Ser Glu Ala Lys Leu Glu Ser Phe Pro Ser
440 445 450

Ile Gly Pro Gln Arg Leu Ser Phe Asp Ser Ala Thr Phe Met Glu
455 460 465

Ser Glu Lys Gln Asp Val His Glu Phe Leu Leu Glu Asn Leu Thr
470 475 480

Asn Gly Gly Ile Leu Glu Leu Met Met Arg Tyr Leu Lys Ala Met
485 490 495

Gly His Lys Phe Leu Val Arg Trp Pro Pro Gly Leu Ala Glu Val
500 505 510

Val Leu Ser Val Tyr His Ser Trp Arg Arg His Ser Thr Ser Leu
515 520 525

Pro Asn Pro Leu Leu Arg Asp Cys Ser Asn Lys His Ile Lys Asp
530 535 540

Arg Met Leu Met Ser Leu Ser Cys Met Glu Leu Gln Leu Asp Gln
545 550 555

Trp Leu Leu Thr Lys Gly Arg Ser Ser Ala Val Ser Pro Arg Asn
560 565 570

Cys Pro Ala Gly Met Val Asn Gly Arg Phe Gly Pro Asp Phe Pro
575 580 585

Gly Thr His Cys Leu Gly Asp Leu Leu Gln Leu Ser Phe Ala Ser
590 595 600

Ser Gln Arg Asp Leu Phe Glu Asp Gly Trp Leu Glu Phe Val Val
605 610 615

Arg Val Tyr Trp Leu Lys Ala Arg Phe Leu Ala Leu Gln Gly Asp
620 625 630

Met Glu Gln Ala Leu Glu Asn Tyr Asp Ile Cys Thr Glu Met Leu
635 640 645

Gln Ser Ser Thr Ala Ile Gln Val Glu Ala Gly Ala Glu Arg Arg
650 655 660

Asp Ile Val Ile Arg Leu Pro Asn Leu His Asn Asp Ser Val Val
665 670 675

Ser Leu Glu Glu Ile Asp Lys Asn Leu Lys Ser Leu Glu Arg Cys
680 685 690

Gln Ser Leu Glu Glu Ile Gln Arg Leu Tyr Glu Ala Gly Asp Tyr
695 700 705

Lys Ala Val Val His Leu Leu Arg Pro Thr Leu Cys Thr Ser Gly
710 715 720

Phe Asp Arg Ala Lys His Leu Glu Phe Met Thr Ser Ile Pro Glu
725 730 735

Arg Pro Ala Gln Leu Leu Leu Leu Gln Asp Ser Leu Leu Arg Leu
740 745 750

Lys Asp Tyr Arg Gln Cys Phe Glu Cys Ser Asp Val Ala Leu Asn
755 760 765

Glu Ala Val Gln Gln Met Val Asn Ser Gly Glu Ala Ala Ala Lys
770 775 780

Glu Glu Trp Val Ala Thr Val Thr Gln Leu Leu Met Gly Ile Glu
785 790 795

Gln Ala Leu Ser Ala Asp Ser Ser Gly Ser Ile Leu Lys Val Ser
800 805 810

Ser Ser Thr Thr Gly Leu Val Arg Leu Thr Asn Asn Leu Ile Gln
815 820 825

Val Ile Asp Cys Ser Met Ala Val Gln Glu Glu Ala Lys Glu Pro
830 835 840

His Val Ser Ser Val Leu Pro Trp Ile Ile Leu His Arg Ile Ile
845 850 855

Trp Gln Glu Glu Asp Thr Phe His Ser Leu Cys His Gln Gln Gln
860 865 870
Leu Gln Asn Pro Ala Glu Glu Gly Met Ser Glu Thr Pro Met Leu
875 880 885
Pro Ser Ser Leu Met Leu Leu Asn Thr Ala His Glu Tyr Leu Gly
890 895 900
Arg Arg Ser Trp Cys Cys Asn Ser Asp Gly Ala Leu Leu Arg Phe
905 910 915
Tyr Val Arg Val Leu Gln Lys Glu Leu Ala Ala Ser Thr Ser Glu
920 925 930
Asp Thr His Pro Tyr Lys Glu Glu Leu Glu Thr Ala Leu Glu Gln
935 940 945
Cys Phe Tyr Cys Leu Tyr Ser Phe Pro Ser Lys Lys Ser Lys Ala
950 955 960
Arg Tyr Leu Glu Glu His Ser Ala Gln Gln Val Asp Leu Ile Trp
965 970 975
Glu Asp Ala Leu Phe Met Phe Glu Tyr Phe Lys Pro Lys Thr Leu
980 985 990
Pro Glu Phe Asp Ser Tyr Lys Thr Ser Thr Val Ser Ala Asp Leu
995 1000 1005
Ala Asn Leu Leu Lys Arg Ile Ala Thr Ile Val Pro Arg Thr Glu
1010 1015 1020
Arg Pro Ala Leu Ser Leu Asp Lys Val Ser Ala Tyr Ile Glu Gly
1025 1030 1035
Thr Ser Thr Glu Val Pro Cys Leu Pro Glu Gly Ala Asp Pro Ser
1040 1045 1050
Pro Pro Val Val Asn Glu Leu Tyr Tyr Leu Leu Ala Asp Tyr His
1055 1060 1065
Phe Lys Asn Lys Glu Gln Ser Lys Ala Ile Lys Phe Tyr Met His
1070 1075 1080
Asp Ile Cys Ile Cys Pro Asn Arg Phe Asp Ser Trp Ala Gly Met
1085 1090 1095
Ala Leu Ala Arg Ala Ser Arg Ile Gln Asp Lys Leu Asn Ser Asn
1100 1105 1110
Glu Leu Lys Ser Asp Gly Pro Ile Trp Lys His Ala Thr Pro Val
1115 1120 1125
Leu Asn Cys Phe Arg Arg Ala Leu Glu Ile Asp Ser Ser Asn Leu
1130 1135 1140
Ser Leu Trp Ile Glu Tyr Gly Thr Met Ser Tyr Ala Leu His Ser
1145 1150 1155
Phe Ala Ser Arg Gln Leu Lys Gln Trp Arg Gly Glu Leu Pro Pro
1160 1165 1170
Glu Leu Val Gln Gln Met Glu Gly Arg Arg Asp Ser Met Leu Glu
1175 1180 1185
Thr Ala Lys His Cys Phe Thr Ser Ala Ala Arg Cys Glu Gly Asp
1190 1195 1200
Gly Asp Glu Glu Glu Trp Leu Ile His Tyr Met Leu Gly Lys Val
1205 1210 1215
Ala Glu Lys Gln Gln Gln Pro Pro Thr Val Tyr Leu Leu His Tyr
1220 1225 1230
Arg Gln Ala Gly His Tyr Leu His Glu Glu Ala Ala Arg Tyr Pro
1235 1240 1245
Lys Lys Ile His Tyr His Asn Pro Pro Glu Leu Ala Met Glu Ala
1250 1255 1260
Leu Glu Val Tyr Phe Arg Leu His Ala Ser Ile Leu Lys Leu Leu
1265 1270 1275

```
Gly Lys Pro Asp Ser Gly Val Gly Ala Glu Val Leu Val Asn Phe
                1280                1285                1290
Met Lys Glu Ala Ala Glu Gly Pro Phe Ala Arg Gly Glu Glu Lys
                1295                1300                1305
Asn Thr Pro Lys Ala Ser Glu Lys Glu Lys Ala Cys Leu Val Asp
                1310                1315                1320
Glu Asp Ser His Ser Ser Ala Gly Thr Leu Pro Gly Pro Gly Ala
                1325                1330                1335
Ser Leu Pro Ser Ser Ser Gly Pro Gly Leu Thr Ser Pro Pro Tyr
                1340                1345                1350
Thr Ala Thr Pro Ile Asp His Asp Tyr Val Lys Cys Lys Lys Pro
                1355                1360                1365
His Gln Gln Ala Thr Pro Asp Asp Arg Ser Gln Asp Ser Thr Ala
                1370                1375                1380
Val Ala Leu Ser Asp Ser Ser Ser Thr Gln Asp Phe Phe Asn Glu
                1385                1390                1395
Pro Thr Ser Leu Leu Glu Gly Ser Arg Lys Ser Tyr Thr Glu Lys
                1400                1405                1410
Arg Leu Pro Ile Leu Ser Ser Gln Ala Gly Ala Thr Gly Lys Asp
                1415                1420                1425
Leu Gln Gly Ala Thr Glu Glu Arg Gly Lys Asn Glu Glu Ser Leu
                1430                1435                1440
Glu Ser Thr Glu Gly Phe Arg Ala Ala Glu Gln Gly Val Gln Lys
                1445                1450                1455
Pro Ala Ala Glu Thr Pro Ala Ser Ala Cys Ile Pro Gly Lys Pro
                1460                1465                1470
Ser Ala Ser Thr Pro Thr Leu Trp Asp Gly Lys Lys Arg Gly Asp
                1475                1480                1485
Leu Pro Gly Glu Pro Val Ala Phe Pro Gln Gly Leu Pro Ala Gly
                1490                1495                1500
Ala Glu Glu Gln Arg Gln Phe Leu Thr Glu Gln Cys Ile Ala Ser
                1505                1510                1515
Phe Arg Leu Cys Leu Ser Arg Phe Pro Gln His Tyr Lys Ser Leu
                1520                1525                1530
Tyr Arg Leu Ala Phe Leu Tyr Thr Tyr Ser Lys Thr His Arg Asn
                1535                1540                1545
Leu Gln Trp Ala Arg Asp Val Leu Leu Gly Ser Ser Ile Pro Trp
                1550                1555                1560
Gln Gln Leu Gln His Met Pro Ala Gln Gly Leu Phe Cys Glu Arg
                1565                1570                1575
Asn Lys Thr Asn Phe Phe Asn Gly Ile Trp Arg Ile Pro Val Asp
                1580                1585                1590
Glu Ile Asp Arg Pro Gly Ser Phe Ala Trp His Met Asn Arg Ser
                1595                1600                1605
Ile Val Leu Leu Leu Lys Val Leu Ala Gln Leu Arg Asp His Ser
                1610                1615                1620
Thr Leu Leu Lys Val Ser Ser Met Leu Gln Arg Thr Pro Asp Gln
                1625                1630                1635
Gly Lys Lys Tyr Leu Arg Asp Ala Asp Arg Gln Val Leu Ala Gln
                1640                1645                1650
Arg Ala Phe Ile Leu Thr Val Lys Val Leu Glu Asp Thr Leu Ser
                1655                1660                1665
Glu Leu Ala Glu Gly Ser Glu Arg Pro Gly Pro Lys Val Cys Gly
                1670                1675                1680
Leu Pro Gly Ala Arg Met Thr Thr Asp Val Ser His Lys Ala Ser
                1685                1690                1695
```

```
Pro Glu Asp Gly Gln Glu Gly Leu Pro Gln Pro Lys Lys Pro Pro
                1700                1705                1710
Leu Ala Asp Gly Ser Gly Pro Gly Pro Glu Pro Gly Gly Lys Val
                1715                1720                1725
Gly Leu Leu Asn His Arg Pro Val Ala Met Asp Ala Gly Asp Ser
                1730                1735                1740
Ala Asp Gln Ser Gly Glu Arg Lys Asp Lys Glu Ser Pro Arg Ala
                1745                1750                1755
Gly Pro Thr Glu Pro Met Asp Thr Ser Glu Ala Thr Val Cys His
                1760                1765                1770
Ser Asp Leu Glu Arg Thr Pro Pro Leu Leu Pro Gly Arg Pro Ala
                1775                1780                1785
Arg Asp Arg Gly Pro Glu Ser Arg Pro Thr Glu Leu Ser Leu Glu
                1790                1795                1800
Glu Leu Ser Ile Ser Ala Arg Gln Gln Pro Thr Pro Leu Thr Pro
                1805                1810                1815
Ala Gln Pro Ala Pro Ala Pro Ala Pro Ala Thr Thr Thr Gly Thr
                1820                1825                1830
Arg Ala Gly Gly His Pro Glu Glu Pro Leu Ser Arg Leu Ser Arg
                1835                1840                1845
Lys Arg Lys Leu Leu Glu Asp Thr Glu Ser Gly Lys Thr Leu Leu
                1850                1855                1860
Leu Asp Ala Tyr Arg Val Trp Gln Gln Gly Gln Lys Gly Val Ala
                1865                1870                1875
Tyr Asp Leu Gly Arg Val Glu Arg Ile Met Ser Glu Thr Tyr Met
                1880                1885                1890
Leu Ile Lys Gln Val Asp Glu Glu Ala Ala Leu Glu Gln Ala Val
                1895                1900                1905
Lys Phe Cys Gln Val His Leu Gly Ala Ala Ala Gln Arg Gln Ala
                1910                1915                1920
Ser Gly Asp Thr Pro Thr Thr Pro Lys His Pro Lys Asp Ser Arg
                1925                1930                1935
Glu Asn Phe Phe Pro Val Thr Val Val Pro Thr Ala Pro Asp Pro
                1940                1945                1950
Val Pro Ala Asp Ser Val Gln Arg Pro Ser Asp Ala His Thr Lys
                1955                1960                1965
Pro Arg Pro Ala Leu Ala Ala Ala Thr Thr Ile Ile Thr Cys Pro
                1970                1975                1980
Pro Ser Ala Ser Ala Ser Thr Leu Asp Gln Ser Lys Asp Pro Gly
                1985                1990                1995
Pro Pro Arg Pro His Arg Pro Glu Ala Thr Pro Ser Met Ala Ser
                2000                2005                2010
Leu Gly Pro Glu Gly Glu Glu Leu Ala Arg Val Ala Glu Gly Thr
                2015                2020                2025
Ser Phe Pro Pro Gln Glu Pro Arg His Ser Pro Gln Val Lys Met
                2030                2035                2040
Ala Pro Thr Ser Ser Pro Ala Glu Pro His Cys Trp Pro Ala Glu
                2045                2050                2055
Ala Ala Leu Gly Thr Gly Ala Glu Pro Thr Cys Ser Gln Ala Ala
                2060                2065                2070
Ser Ser Lys Ala Pro Ser Ser Gly Ser Ala Gln Pro Pro Glu Gly
                2075                2080                2085
His Pro Gly Lys Pro Glu Pro Ser Arg Ala Lys Ser Arg Pro Leu
                2090                2095                2100
Pro Asn Met Pro Lys Leu Val Ile Pro Ser Ala Ala Thr Lys Phe
                2105                2110                2115
```

```
Pro Pro Glu Ile Thr Val Thr Pro Pro Thr Pro Thr Leu Leu Ser
                2120                2125                2130
Pro Lys Gly Ser Ile Ser Glu Glu Thr Lys Gln Lys Leu Lys Ser
                2135                2140                2145
Ala Ile Leu Ser Ala Gln Ser Ala Ala Asn Val Arg Lys Glu Ser
                2150                2155                2160
Leu Cys Gln Pro Ala Leu Glu Val Leu Glu Thr Ser Ser Gln Glu
                2165                2170                2175
Ser Ser Leu Glu Ser Glu Thr Asp Glu Asp Asp Asp Tyr Met Asp
                2180                2185                2190
Ile


<210> 11
<211> 347
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90086258CD1

<400> 11
Met Val Ser Ser Gln Lys Leu Glu Lys Pro Ile Glu Met Gly Ser
  1               5                  10                  15
Ser Glu Pro Leu Pro Ile Ala Asp Gly Gly Arg Arg Arg Lys Lys
                 20                  25                  30
Lys Arg Arg Gly Arg Ala Thr Asp Ser Leu Pro Gly Lys Phe Glu
                 35                  40                  45
Asp Met Tyr Lys Leu Thr Ser Glu Leu Leu Gly Glu Gly Ala Tyr
                 50                  55                  60
Ala Lys Val Gln Gly Ala Val Ser Leu Gln Asn Gly Lys Glu Tyr
                 65                  70                  75
Ala Val Lys Ile Ile Glu Lys Gln Ala Gly His Ser Arg Ser Arg
                 80                  85                  90
Val Phe Arg Glu Val Glu Thr Leu Tyr Gln Cys Gln Gly Asn Lys
                 95                 100                 105
Asn Ile Leu Glu Leu Ile Glu Phe Phe Glu Asp Asp Thr Arg Phe
                110                 115                 120
Tyr Leu Val Phe Glu Lys Leu Gln Gly Gly Ser Ile Leu Ala His
                125                 130                 135
Ile Gln Lys Gln Lys His Phe Asn Glu Arg Glu Ala Ser Arg Val
                140                 145                 150
Val Arg Asp Val Ala Ala Ala Leu Asp Phe Leu His Thr Lys Gly
                155                 160                 165
Ile Ala His Arg Asp Leu Lys Pro Glu Asn Ile Leu Cys Glu Ser
                170                 175                 180
Pro Glu Lys Val Ser Pro Val Lys Ile Cys Asp Phe Asp Leu Gly
                185                 190                 195
Ser Gly Met Lys Leu Asn Asn Ser Cys Thr Pro Ile Thr Thr Pro
                200                 205                 210
Glu Leu Thr Thr Pro Cys Gly Ser Ala Glu Tyr Met Ala Pro Glu
                215                 220                 225
Val Val Glu Val Phe Thr Asp Gln Ala Thr Phe Tyr Asp Lys Arg
                230                 235                 240
Cys Asp Leu Trp Ser Leu Gly Val Val Leu Tyr Ile Met Leu Ser
```

```
                245                 250                 255
Gly Tyr Pro Pro Phe Val Gly His Cys Gly Ala Asp Cys Gly Trp
                260                 265                 270
Asp Arg Gly Glu Val Cys Arg Val Cys Gln Asn Lys Leu Phe Glu
                275                 280                 285
Ser Ile Gln Glu Gly Lys Tyr Glu Phe Pro Asp Lys Asp Trp Ala
                290                 295                 300
His Ile Ser Ser Glu Ala Lys Asp Leu Ile Ser Lys Leu Leu Val
                305                 310                 315
Arg Asp Ala Lys Gln Arg Leu Ser Ala Ala Gln Val Leu Gln His
                320                 325                 330
Pro Trp Val Gln Gly Glu Gln Gln His Asn Gly Pro Asp Ala Leu
                335                 340                 345
Arg Ser


<210> 12
<211> 498
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1967990CD1

<400> 12
Met Pro His Pro Arg Arg Tyr His Ser Ser Glu Arg Gly Ser Arg
  1              5                  10                  15
Gly Ser Tyr Arg Glu His Tyr Arg Ser Arg Lys His Lys Arg Arg
                 20                  25                  30
Arg Ser Arg Ser Trp Ser Ser Ser Ser Asp Arg Thr Arg Arg Arg
                 35                  40                  45
Arg Arg Glu Asp Ser Tyr His Val Arg Ser Arg Ser Tyr Asp Asp
                 50                  55                  60
Arg Ser Ser Asp Arg Arg Val Tyr Asp Arg Arg Tyr Cys Gly Ser
                 65                  70                  75
Tyr Arg Arg Asn Asp Tyr Ser Arg Asp Arg Gly Asp Ala Tyr Tyr
                 80                  85                  90
Asp Thr Asp Tyr Arg His Ser Tyr Glu Tyr Gln Arg Glu Asn Ser
                 95                 100                 105
Ser Tyr Arg Ser Gln Arg Ser Ser Arg Arg Lys His Arg Arg Arg
                110                 115                 120
Arg Arg Arg Ser Arg Thr Phe Ser Arg Ser Ser Ser Gln His Ser
                125                 130                 135
Ser Arg Arg Ala Lys Ser Val Glu Asp Asp Ala Glu Gly His Leu
                140                 145                 150
Ile Tyr His Val Gly Asp Trp Leu Gln Glu Arg Tyr Glu Ile Val
                155                 160                 165
Ser Thr Leu Gly Glu Gly Thr Phe Gly Arg Val Val Gln Cys Val
                170                 175                 180
Asp His Arg Arg Gly Gly Ala Arg Val Ala Leu Lys Ile Ile Lys
                185                 190                 195
Asn Val Glu Lys Tyr Lys Glu Ala Ala Arg Leu Glu Ile Asn Val
                200                 205                 210
Leu Glu Lys Ile Asn Glu Lys Asp Pro Asp Asn Lys Asn Leu Cys
                215                 220                 225
```

```
Val Gln Met Phe Asp Trp Phe Asp Tyr His Gly His Met Cys Ile
                230                 235                 240
Ser Phe Glu Leu Leu Gly Leu Ser Thr Phe Asp Phe Leu Lys Asp
                245                 250                 255
Asn Asn Tyr Leu Pro Tyr Pro Ile His Gln Val Arg His Met Ala
                260                 265                 270
Phe Gln Leu Cys Gln Ala Val Lys Phe Leu His Asp Asn Lys Leu
                275                 280                 285
Thr His Thr Asp Leu Lys Pro Glu Asn Ile Leu Phe Val Asn Ser
                290                 295                 300
Asp Tyr Glu Leu Thr Tyr Asn Leu Glu Lys Lys Arg Asp Glu Arg
                305                 310                 315
Ser Val Lys Ser Thr Ala Val Arg Val Val Asp Phe Gly Ser Ala
                320                 325                 330
Thr Phe Asp His Glu His His Ser Thr Ile Val Ser Thr Arg His
                335                 340                 345
Tyr Arg Ala Pro Glu Val Ile Leu Glu Leu Gly Trp Ser Gln Pro
                350                 355                 360
Cys Asp Val Trp Ser Ile Gly Cys Ile Ile Phe Glu Tyr Tyr Val
                365                 370                 375
Gly Phe Thr Leu Phe Gln Thr His Asp Asn Arg Glu His Leu Ala
                380                 385                 390
Met Met Glu Arg Ile Leu Gly Pro Ile Pro Ser Arg Met Ile Arg
                395                 400                 405
Lys Thr Arg Lys Gln Lys Tyr Phe Tyr Arg Gly Arg Leu Asp Trp
                410                 415                 420
Asp Glu Asn Thr Ser Ala Gly Arg Tyr Val Arg Glu Asn Cys Lys
                425                 430                 435
Pro Leu Arg Arg Tyr Leu Thr Ser Glu Ala Glu Glu His His Gln
                440                 445                 450
Leu Phe Asp Leu Ile Glu Ser Met Leu Glu Tyr Glu Pro Ala Lys
                455                 460                 465
Arg Leu Thr Leu Gly Glu Ala Leu Gln His Pro Phe Phe Ala Arg
                470                 475                 480
Leu Arg Ala Glu Pro Pro Asn Lys Leu Trp Asp Ser Ser Arg Asp
                485                 490                 495
Ile Ser Arg
```

<210> 13
<211> 1081
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3810039CD1

<400> 13

```
Met Glu Pro Ser Arg Ala Leu Leu Gly Cys Leu Ala Ser Ala Ala
  1               5                  10                  15
Ala Ala Ala Pro Pro Gly Glu Asp Gly Ala Gly Ala Gly Ala Glu
                 20                  25                  30
Glu Glu Glu Glu Glu Glu Glu Glu Ala Ala Ala Ala Val Gly Pro
                 35                  40                  45
Gly Glu Leu Gly Cys Asp Ala Pro Leu Pro Tyr Trp Thr Ala Val
```

```
                50                  55                  60
Phe Glu Tyr Glu Ala Ala Gly Glu Asp Glu Leu Thr Leu Arg Leu
                65                  70                  75
Gly Asp Val Val Glu Val Leu Ser Lys Asp Ser Gln Val Ser Gly
                80                  85                  90
Asp Glu Gly Trp Trp Thr Gly Gln Leu Asn Gln Arg Val Gly Ile
                95                  100                 105
Phe Pro Ser Asn Tyr Val Thr Pro Arg Ser Ala Phe Ser Ser Arg
                110                 115                 120
Cys Gln Pro Gly Gly Glu Asp Pro Ser Cys Tyr Pro Pro Ile Gln
                125                 130                 135
Leu Leu Glu Ile Asp Phe Ala Glu Leu Thr Leu Glu Glu Ile Ile
                140                 145                 150
Gly Ile Gly Gly Phe Gly Lys Val Tyr Arg Ala Phe Trp Ile Gly
                155                 160                 165
Asp Glu Val Ala Val Lys Ala Ala Arg His Asp Pro Asp Glu Asp
                170                 175                 180
Ile Ser Gln Thr Ile Glu Asn Val Arg Gln Glu Ala Lys Leu Phe
                185                 190                 195
Ala Met Leu Lys His Pro Asn Ile Ile Ala Leu Arg Gly Val Cys
                200                 205                 210
Leu Lys Glu Pro Asn Leu Cys Leu Val Met Glu Phe Ala Arg Gly
                215                 220                 225
Gly Pro Leu Asn Arg Val Leu Ser Gly Lys Arg Ile Pro Pro Asp
                230                 235                 240
Ile Leu Val Asn Trp Ala Val Gln Ile Ala Arg Gly Met Asn Tyr
                245                 250                 255
Leu His Asp Glu Ala Ile Val Pro Ile Ile His Arg Asp Leu Lys
                260                 265                 270
Ser Ser Asn Ile Leu Ile Leu Gln Lys Val Glu Asn Gly Asp Leu
                275                 280                 285
Ser Asn Lys Ile Leu Lys Ile Thr Asp Phe Gly Leu Ala Arg Glu
                290                 295                 300
Trp His Arg Thr Thr Lys Met Ser Ala Ala Gly Thr Tyr Ala Trp
                305                 310                 315
Met Ala Pro Glu Val Ile Arg Ala Ser Met Phe Ser Lys Gly Ser
                320                 325                 330
Asp Val Trp Ser Tyr Gly Val Leu Leu Trp Glu Leu Leu Thr Gly
                335                 340                 345
Glu Val Pro Phe Arg Gly Ile Asp Gly Leu Ala Val Ala Tyr Gly
                350                 355                 360
Val Ala Met Asn Lys Leu Ala Leu Pro Ile Pro Ser Thr Cys Pro
                365                 370                 375
Glu Pro Phe Ala Lys Leu Met Glu Asp Cys Trp Asn Pro Asp Pro
                380                 385                 390
His Ser Arg Pro Ser Phe Thr Asn Ile Leu Asp Gln Leu Thr Thr
                395                 400                 405
Ile Glu Glu Ser Gly Phe Phe Glu Met Pro Lys Asp Ser Phe His
                410                 415                 420
Cys Leu Gln Asp Asn Trp Lys His Glu Ile Gln Glu Met Phe Asp
                425                 430                 435
Gln Leu Arg Ala Lys Glu Lys Glu Leu Arg Thr Trp Glu Glu Glu
                440                 445                 450
Leu Thr Arg Ala Ala Leu Gln Gln Lys Asn Gln Glu Glu Leu Leu
                455                 460                 465
Arg Arg Arg Glu Gln Glu Leu Ala Glu Arg Glu Ile Asp Ile Leu
```

```
                470                 475                 480
Glu Arg Glu Leu Asn Ile Ile Ile His Gln Leu Cys Gln Glu Lys
                485                 490                 495
Pro Arg Val Lys Lys Arg Lys Gly Lys Phe Arg Lys Ser Arg Leu
                500                 505                 510
Lys Leu Lys Asp Gly Asn Arg Ile Ser Leu Pro Ser Asp Phe Gln
                515                 520                 525
His Lys Phe Thr Val Gln Ala Ser Pro Thr Met Asp Lys Arg Lys
                530                 535                 540
Ser Leu Ile Asn Ser Arg Ser Ser Pro Pro Ala Ser Pro Thr Ile
                545                 550                 555
Ile Pro Arg Leu Arg Ala Ile Gln Leu Thr Pro Gly Glu Ser Ser
                560                 565                 570
Lys Thr Trp Gly Arg Ser Ser Val Val Pro Lys Glu Glu Gly Glu
                575                 580                 585
Glu Glu Glu Lys Arg Ala Pro Lys Lys Lys Gly Arg Thr Trp Gly
                590                 595                 600
Pro Gly Thr Leu Gly Gln Lys Glu Leu Ala Ser Gly Asp Glu Gly
                605                 610                 615
Leu Lys Ser Leu Val Asp Gly Tyr Lys Gln Trp Ser Ser Ser Ala
                620                 625                 630
Pro Asn Leu Val Lys Gly Pro Arg Ser Ser Pro Ala Leu Pro Gly
                635                 640                 645
Phe Thr Ser Leu Met Glu Met Glu Asp Glu Asp Ser Glu Gly Pro
                650                 655                 660
Gly Ser Gly Glu Ser Arg Leu Gln His Ser Pro Ser Gln Ser Tyr
                665                 670                 675
Leu Cys Ile Pro Phe Pro Arg Gly Glu Asp Gly Asp Gly Pro Ser
                680                 685                 690
Ser Asp Gly Ile His Glu Glu Pro Thr Pro Val Asn Ser Ala Thr
                695                 700                 705
Ser Thr Pro Gln Leu Thr Pro Thr Asn Ser Leu Lys Arg Gly Gly
                710                 715                 720
Ala His His Arg Arg Cys Glu Val Ala Leu Leu Gly Cys Gly Ala
                725                 730                 735
Val Leu Ala Ala Thr Gly Leu Gly Phe Asp Leu Leu Glu Ala Gly
                740                 745                 750
Lys Cys Gln Leu Leu Pro Leu Glu Glu Pro Glu Pro Pro Ala Arg
                755                 760                 765
Glu Glu Lys Lys Arg Arg Glu Gly Leu Phe Gln Arg Ser Ser Arg
                770                 775                 780
Pro Arg Arg Ser Thr Ser Pro Pro Ser Arg Lys Leu Phe Lys Lys
                785                 790                 795
Glu Glu Pro Met Leu Leu Leu Gly Asp Pro Ser Ala Ser Leu Thr
                800                 805                 810
Leu Leu Ser Leu Ser Ser Ile Ser Glu Cys Asn Ser Thr Arg Ser
                815                 820                 825
Leu Leu Arg Ser Asp Ser Asp Glu Ile Val Val Tyr Glu Met Pro
                830                 835                 840
Val Ser Pro Val Glu Ala Pro Pro Leu Ser Pro Cys Thr His Asn
                845                 850                 855
Pro Leu Val Asn Val Arg Val Glu Arg Phe Lys Arg Asp Pro Asn
                860                 865                 870
Gln Ser Leu Thr Pro Thr His Val Thr Leu Thr Thr Pro Ser Gln
                875                 880                 885
Pro Ser Ser His Arg Arg Thr Pro Ser Asp Gly Ala Leu Lys Pro
```

```
                890                 895                 900
Glu Thr Leu Leu Ala Ser Arg Ser Pro Ser Ser Asn Gly Leu Ser
                905                 910                 915
Pro Ser Pro Gly Ala Gly Met Leu Lys Thr Pro Ser Pro Ser Arg
                920                 925                 930
Asp Pro Gly Glu Phe Pro Arg Leu Pro Asp Pro Asn Val Val Phe
                935                 940                 945
Pro Pro Thr Pro Arg Arg Trp Asn Thr Gln Gln Asp Ser Thr Leu
                950                 955                 960
Glu Arg Pro Lys Thr Leu Glu Phe Leu Pro Arg Pro Arg Pro Ser
                965                 970                 975
Ala Asn Arg Gln Arg Leu Asp Pro Trp Trp Phe Val Ser Pro Ser
                980                 985                 990
His Ala Arg Ser Thr Ser Pro Ala Asn Ser Ser Ser Thr Glu Thr
                995                 1000                1005
Pro Ser Asn Leu Asp Ser Cys Phe Ala Ser Ser Ser Ser Thr Val
                1010                1015                1020
Glu Glu Arg Pro Gly Leu Pro Ala Leu Leu Pro Phe Gln Ala Gly
                1025                1030                1035
Pro Leu Pro Pro Thr Glu Arg Thr Leu Leu Asp Leu Asp Ala Glu
                1040                1045                1050
Gly Gln Ser Gln Asp Ser Thr Val Pro Leu Cys Arg Ala Glu Leu
                1055                1060                1065
Asn Thr His Arg Pro Ala Pro Tyr Glu Ile Gln Gln Glu Phe Trp
                1070                1075                1080
Ser


<210> 14
<211> 170
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8032337CD1

<400> 14
Met Ala Ser Glu Ala Glu Lys Thr Phe His Arg Phe Ala Ala Phe
  1               5                  10                  15
Gly Glu Ser Ser Ser Ser Gly Thr Glu Met Asn Asn Lys Asn Phe
                 20                  25                  30
Ser Lys Leu Cys Lys Asp Cys Gly Ile Met Asp Gly Lys Thr Val
                 35                  40                  45
Thr Ser Thr Asp Val Asp Ile Val Phe Ser Lys Val Lys Ala Lys
                 50                  55                  60
Asn Ala Arg Thr Ile Thr Phe Gln Gln Phe Lys Glu Ala Val Lys
                 65                  70                  75
Glu Leu Gly Gln Lys Arg Phe Lys Gly Lys Ser Pro Asp Glu Val
                 80                  85                  90
Leu Glu Asn Ile Tyr Gly Leu Met Glu Gly Lys Asp Pro Ala Thr
                 95                 100                 105
Thr Gly Ala Thr Lys Ala Thr Thr Val Gly Ala Val Asp Arg Leu
                110                 115                 120
Thr Asp Thr Ser Lys Tyr Thr Gly Thr His Lys Glu Arg Phe Asp
                125                 130                 135
```

```
Glu Ser Gly Lys Gly Lys Gly Ile Ala Gly Arg Glu Glu Met Thr
                140                 145                 150
Asp Asn Thr Gly Tyr Val Ser Gly Tyr Lys Gly Ser Gly Thr Tyr
                155                 160                 165
Asp Lys Lys Thr Lys
                170

<210> 15
<211> 893
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506411CD1

<400> 15
Met Pro Val Arg Lys Gln Asp Thr Gln Arg Ala Leu His Leu Leu
 1                5                   10                  15
Glu Glu Tyr Arg Ser Lys Leu Ser Gln Thr Glu Asp Arg Gln Leu
                 20                  25                  30
Arg Ser Ser Ile Glu Arg Val Ile Asn Ile Phe Gln Ser Asn Leu
                 35                  40                  45
Phe Gln Ala Leu Ile Asp Ile Gln Glu Phe Tyr Glu Val Thr Leu
                 50                  55                  60
Leu Asp Asn Pro Lys Cys Ile Asp Arg Ser Lys Pro Ser Glu Pro
                 65                  70                  75
Ile Gln Pro Val Asn Thr Trp Glu Ile Ser Ser Leu Pro Ser Ser
                 80                  85                  90
Thr Val Thr Ser Glu Thr Leu Pro Ser Ser Leu Ser Pro Ser Val
                 95                 100                 105
Glu Lys Tyr Arg Tyr Gln Asp Glu Asp Thr Pro Pro Gln Glu His
                110                 115                 120
Ile Ser Pro Gln Ile Thr Asn Glu Val Ile Gly Pro Glu Leu Val
                125                 130                 135
His Val Ser Glu Lys Asn Leu Ser Glu Ile Glu Asn Val His Gly
                140                 145                 150
Phe Val Ser His Ser His Ile Ser Pro Ile Lys Ala Asn Pro Pro
                155                 160                 165
Pro Val Leu Val Asn Thr Asp Ser Leu Glu Thr Pro Thr Tyr Val
                170                 175                 180
Asn Gly Thr Asp Ala Asp Tyr Glu Tyr Glu Glu Ile Thr Leu Glu
                185                 190                 195
Arg Gly Asn Ser Gly Leu Gly Phe Ser Ile Ala Gly Gly Thr Asp
                200                 205                 210
Asn Pro His Ile Gly Asp Asp Ser Ser Ile Phe Ile Thr Lys Ile
                215                 220                 225
Ile Thr Gly Gly Ala Ala Ala Gln Asp Gly Arg Leu Arg Val Asn
                230                 235                 240
Asp Cys Ile Leu Arg Val Asn Glu Val Asp Val Arg Asp Val Thr
                245                 250                 255
His Ser Lys Ala Val Glu Ala Leu Lys Glu Ala Gly Ser Ile Val
                260                 265                 270
Arg Leu Tyr Val Lys Arg Arg Lys Pro Val Ser Glu Lys Ile Met
                275                 280                 285
Glu Ile Lys Leu Ile Lys Gly Pro Lys Gly Leu Gly Phe Ser Ile
```

```
                    290                         295                         300
Ala Gly Gly Val Gly Asn Gln His Ile Pro Gly Asp Asn Ser Ile
                    305                         310                         315
Tyr Val Thr Lys Ile Ile Glu Gly Gly Ala Ala His Lys Asp Gly
                    320                         325                         330
Lys Leu Gln Ile Gly Asp Lys Leu Leu Ala Val Asn Asn Val Cys
                    335                         340                         345
Leu Glu Glu Val Thr His Glu Glu Ala Val Thr Ala Leu Lys Asn
                    350                         355                         360
Thr Ser Asp Phe Val Tyr Leu Lys Val Ala Lys Pro Thr Ser Met
                    365                         370                         375
Tyr Met Asn Asp Gly Tyr Ala Pro Pro Asp Ile Thr Asn Ser Ser
                    380                         385                         390
Ser Gln Pro Val Asp Asn His Val Ser Pro Ser Ser Phe Leu Gly
                    395                         400                         405
Gln Thr Pro Ala Ser Pro Ala Arg Tyr Ser Pro Val Ser Lys Ala
                    410                         415                         420
Val Leu Gly Asp Asp Glu Ile Thr Arg Glu Pro Arg Lys Val Val
                    425                         430                         435
Leu His Arg Gly Ser Thr Gly Leu Gly Phe Asn Ile Val Gly Gly
                    440                         445                         450
Glu Asp Gly Glu Gly Ile Phe Ile Ser Phe Ile Leu Ala Gly Gly
                    455                         460                         465
Pro Ala Asp Leu Ser Gly Glu Leu Arg Lys Gly Asp Arg Ile Ile
                    470                         475                         480
Ser Val Asn Ser Val Asp Leu Arg Ala Ala Ser His Glu Gln Ala
                    485                         490                         495
Ala Ala Ala Leu Lys Asn Ala Gly Gln Ala Val Thr Ile Val Ala
                    500                         505                         510
Gln Tyr Arg Pro Glu Glu Tyr Ser Arg Phe Glu Ala Lys Ile His
                    515                         520                         525
Asp Leu Arg Glu Gln Met Met Asn Ser Ser Ile Ser Ser Gly Ser
                    530                         535                         540
Gly Ser Leu Arg Thr Ser Gln Lys Arg Ser Leu Tyr Val Arg Ala
                    545                         550                         555
Leu Phe Asp Tyr Asp Lys Thr Lys Asp Ser Gly Leu Pro Ser Gln
                    560                         565                         570
Gly Leu Asn Phe Lys Phe Gly Asp Ile Leu His Val Ile Asn Ala
                    575                         580                         585
Ser Asp Asp Glu Trp Trp Gln Ala Arg Gln Val Thr Pro Asp Gly
                    590                         595                         600
Glu Ser Asp Glu Val Gly Val Ile Pro Ser Lys Arg Arg Val Glu
                    605                         610                         615
Lys Lys Glu Arg Ala Arg Leu Lys Thr Val Lys Phe Asn Ser Lys
                    620                         625                         630
Thr Arg Asp Lys Gly Gln Ser Phe Asn Asp Lys Arg Lys Lys Asn
                    635                         640                         645
Leu Phe Ser Arg Lys Phe Pro Phe Tyr Lys Asn Lys Asp Gln Ser
                    650                         655                         660
Glu Gln Glu Thr Ser Asp Ala Asp Gln His Val Thr Ser Asn Ala
                    665                         670                         675
Ser Asp Ser Glu Ser Ser Tyr Arg Gly Gln Glu Glu Tyr Val Leu
                    680                         685                         690
Ser Tyr Glu Pro Val Asn Gln Gln Glu Val Asn Tyr Thr Arg Pro
                    695                         700                         705
Val Ile Ile Leu Gly Pro Met Lys Asp Arg Ile Asn Asp Asp Leu
```

710 715 720
Ile Ser Glu Phe Pro Asp Lys Phe Gly Ser Cys Val Pro His Thr
725 730 735
Thr Arg Pro Lys Arg Asp Tyr Glu Val Asp Gly Arg Asp Tyr His
740 745 750
Phe Val Thr Ser Arg Glu Gln Met Glu Lys Asp Ile Gln Glu His
755 760 765
Lys Phe Ile Glu Ala Gly Gln Tyr Asn Asn His Leu Tyr Gly Thr
770 775 780
Ser Val Gln Ser Val Arg Glu Val Ala Glu Lys Gly Lys His Cys
785 790 795
Ile Leu Asp Val Ser Gly Asn Ala Ile Lys Arg Leu Gln Ile Ala
800 805 810
Gln Leu Tyr Pro Ile Ser Ile Phe Ile Lys Pro Lys Ser Met Glu
815 820 825
Asn Ile Met Glu Met Asn Lys Arg Leu Thr Glu Glu Gln Ala Arg
830 835 840
Lys Thr Phe Glu Arg Ala Met Lys Leu Glu Gln Glu Phe Thr Glu
845 850 855
His Phe Thr Ala Ile Val Gln Gly Asp Thr Leu Glu Asp Ile Tyr
860 865 870
Asn Gln Val Lys Gln Ile Ile Glu Glu Gln Ser Gly Ser Tyr Ile
875 880 885
Trp Val Pro Ala Lys Glu Lys Leu
890

<210> 16
<211> 1796
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2658834CD1

<400> 16
Met Ala Pro Glu Pro Ala Pro Gly Arg Thr Met Val Pro Leu Val
1 5 10 15
Pro Ala Leu Val Met Leu Gly Leu Val Ala Gly Ala His Gly Asp
20 25 30
Ser Lys Pro Val Phe Ile Lys Val Pro Glu Asp Gln Thr Gly Leu
35 40 45
Ser Gly Gly Val Ala Ser Phe Val Cys Gln Ala Thr Gly Glu Pro
50 55 60
Lys Pro Arg Ile Thr Trp Met Lys Lys Gly Lys Lys Val Ser Ser
65 70 75
Gln Arg Phe Glu Val Ile Glu Phe Asp Asp Gly Ala Gly Ser Val
80 85 90
Leu Arg Ile Gln Pro Leu Arg Val Gln Arg Asp Glu Ala Ile Tyr
95 100 105
Glu Cys Thr Ala Thr Asn Ser Leu Gly Glu Ile Asn Thr Ser Ala
110 115 120
Lys Leu Ser Val Leu Glu Glu Glu Gln Leu Pro Pro Gly Phe Pro
125 130 135
Ser Ile Asp Met Gly Pro Gln Leu Lys Val Val Glu Lys Ala Arg
140 145 150

Thr Ala Thr Met Leu Cys Ala Ala Gly Gly Asn Pro Asp Pro Glu
155 160 165

Ile Ser Trp Phe Lys Asp Phe Leu Pro Val Asp Pro Ala Thr Ser
170 175 180

Asn Gly Arg Ile Lys Gln Leu Arg Ser Gly Ala Leu Gln Ile Glu
185 190 195

Ser Ser Glu Glu Ser Asp Gln Gly Lys Tyr Glu Cys Val Ala Thr
200 205 210

Asn Ser Ala Gly Thr Arg Tyr Ser Ala Pro Ala Asn Leu Tyr Val
215 220 225

Arg Val Arg Arg Val Ala Pro Arg Phe Ser Ile Pro Pro Ser Ser
230 235 240

Gln Glu Val Met Pro Gly Gly Ser Val Asn Leu Thr Cys Val Ala
245 250 255

Val Gly Ala Pro Met Pro Tyr Val Lys Trp Met Met Gly Ala Glu
260 265 270

Glu Leu Thr Lys Glu Asp Glu Met Pro Val Gly Arg Asn Val Leu
275 280 285

Glu Leu Ser Asn Val Val Arg Ser Ala Asn Tyr Thr Cys Val Ala
290 295 300

Ile Ser Ser Leu Gly Met Ile Glu Ala Thr Ala Gln Val Thr Val
305 310 315

Lys Ala Leu Pro Lys Pro Pro Ile Asp Leu Val Val Thr Glu Thr
320 325 330

Thr Ala Thr Ser Val Thr Leu Thr Trp Asp Ser Gly Asn Ser Glu
335 340 345

Pro Val Thr Tyr Tyr Gly Ile Gln Tyr Arg Ala Ala Gly Thr Glu
350 355 360

Gly Pro Phe Gln Glu Val Asp Gly Val Ala Thr Thr Arg Tyr Ser
365 370 375

Ile Gly Gly Leu Ser Pro Phe Ser Glu Tyr Ala Phe Arg Val Leu
380 385 390

Ala Val Asn Ser Ile Gly Arg Gly Pro Pro Ser Glu Ala Val Arg
395 400 405

Ala Arg Thr Gly Glu Gln Ala Pro Ser Ser Pro Pro Arg Arg Val
410 415 420

Gln Ala Arg Met Leu Ser Ala Ser Thr Met Leu Val Gln Trp Glu
425 430 435

Pro Pro Glu Glu Pro Asn Gly Leu Val Arg Gly Tyr Arg Val Tyr
440 445 450

Tyr Thr Pro Asp Ser Arg Arg Pro Pro Asn Ala Trp His Lys His
455 460 465

Asn Thr Asp Ala Gly Leu Leu Thr Thr Val Gly Ser Leu Leu Pro
470 475 480

Gly Ile Thr Tyr Ser Leu Arg Val Leu Ala Phe Thr Ala Val Gly
485 490 495

Asp Gly Pro Pro Ser Pro Thr Ile Gln Val Lys Thr Gln Gln Gly
500 505 510

Val Pro Ala Gln Pro Ala Asp Phe Gln Ala Glu Val Glu Ser Asp
515 520 525

Thr Arg Ile Gln Leu Ser Trp Leu Leu Pro Pro Gln Glu Arg Ile
530 535 540

Ile Met Tyr Glu Leu Val Tyr Trp Ala Ala Glu Asp Glu Asp Gln
545 550 555

Gln His Lys Val Thr Phe Asp Pro Thr Ser Ser Tyr Thr Leu Glu
560 565 570

```
Asp Leu Lys Pro Asp Thr Leu Tyr Arg Phe Gln Leu Ala Ala Arg
                575                 580                 585
Ser Asp Met Gly Val Gly Val Phe Thr Pro Thr Ile Glu Ala Arg
                590                 595                 600
Thr Ala Gln Ser Met Pro Ser Gly Pro Pro Arg Lys Val Glu Val
                605                 610                 615
Glu Pro Leu Asn Ser Thr Ala Val His Val Tyr Trp Lys Leu Pro
                620                 625                 630
Val Pro Ser Lys Gln His Gly Gln Ile Arg Gly Tyr Gln Val Thr
                635                 640                 645
Tyr Val Arg Leu Glu Asn Gly Glu Pro Arg Gly Leu Pro Ile Ile
                650                 655                 660
Gln Asp Val Met Leu Ala Glu Ala Gln Glu Thr Thr Ile Ser Gly
                665                 670                 675
Leu Thr Pro Glu Thr Thr Tyr Ser Val Thr Val Ala Ala Tyr Thr
                680                 685                 690
Thr Lys Gly Asp Gly Ala Arg Ser Lys Pro Lys Ile Val Thr Thr
                695                 700                 705
Thr Gly Ala Val Pro Gly Arg Pro Thr Met Met Ile Ser Thr Thr
                710                 715                 720
Ala Met Asn Thr Ala Leu Leu Gln Trp His Pro Pro Lys Glu Leu
                725                 730                 735
Pro Gly Glu Leu Leu Gly Tyr Arg Leu Gln Tyr Cys Arg Ala Asp
                740                 745                 750
Glu Ala Arg Pro Asn Thr Ile Asp Phe Gly Lys Asp Asp Gln His
                755                 760                 765
Phe Thr Val Thr Gly Leu His Lys Gly Thr Thr Tyr Ile Phe Arg
                770                 775                 780
Leu Ala Ala Lys Asn Arg Ala Gly Leu Gly Glu Glu Phe Glu Lys
                785                 790                 795
Glu Ile Arg Thr Pro Glu Asp Leu Pro Ser Gly Phe Pro Gln Asn
                800                 805                 810
Leu His Val Thr Gly Leu Thr Thr Ser Thr Thr Glu Leu Ala Trp
                815                 820                 825
Asp Pro Pro Val Leu Ala Glu Arg Asn Gly Arg Ile Ile Ser Tyr
                830                 835                 840
Thr Val Val Phe Arg Asp Ile Asn Ser Gln Gln Glu Leu Gln Asn
                845                 850                 855
Ile Thr Thr Asp Thr Arg Phe Thr Leu Thr Gly Leu Lys Pro Asp
                860                 865                 870
Thr Thr Tyr Asp Ile Lys Val Arg Ala Trp Thr Ser Lys Gly Ser
                875                 880                 885
Gly Pro Leu Ser Pro Ser Ile Gln Ser Arg Thr Met Pro Val Glu
                890                 895                 900
Gln Val Phe Ala Lys Asn Phe Arg Val Ala Ala Ala Met Lys Thr
                905                 910                 915
Ser Val Leu Leu Ser Trp Glu Val Pro Asp Ser Tyr Lys Ser Ala
                920                 925                 930
Val Pro Phe Lys Ile Leu Tyr Asn Gly Gln Ser Val Glu Val Asp
                935                 940                 945
Gly His Ser Met Arg Lys Leu Ile Ala Asp Leu Gln Pro Asn Thr
                950                 955                 960
Glu Tyr Ser Phe Val Leu Met Asn Arg Gly Ser Ser Ala Gly Gly
                965                 970                 975
Leu Gln His Leu Val Ser Ile Arg Thr Ala Pro Asp Leu Leu Pro
                980                 985                 990
```

His Lys Pro Leu Pro Ala Ser Ala Tyr Ile Glu Asp Gly Arg Phe
995 1000 1005
Asp Leu Ser Met Pro His Val Gln Asp Pro Ser Leu Val Arg Trp
1010 1015 1020
Phe Tyr Ile Val Val Val Pro Ile Asp Arg Val Gly Gly Ser Met
1025 1030 1035
Leu Thr Pro Arg Trp Ser Thr Pro Glu Glu Leu Glu Leu Asp Glu
1040 1045 1050
Leu Leu Glu Ala Ile Glu Gln Gly Gly Glu Glu Gln Arg Arg Arg
1055 1060 1065
Arg Arg Gln Ala Glu Arg Leu Lys Pro Tyr Val Ala Ala Gln Leu
1070 1075 1080
Asp Val Leu Pro Glu Thr Phe Thr Leu Gly Asp Lys Lys Asn Tyr
1085 1090 1095
Arg Gly Phe Tyr Asn Arg Pro Leu Ser Pro Asp Leu Ser Tyr Gln
1100 1105 1110
Cys Phe Val Leu Ala Ser Leu Lys Glu Pro Met Asp Gln Lys Arg
1115 1120 1125
Tyr Ala Ser Ser Pro Tyr Ser Asp Glu Ile Val Val Gln Val Thr
1130 1135 1140
Pro Ala Gln Gln Gln Glu Glu Pro Glu Met Leu Trp Val Thr Gly
1145 1150 1155
Pro Val Leu Ala Val Ile Leu Ile Ile Leu Ile Val Ile Ala Ile
1160 1165 1170
Leu Leu Phe Lys Arg Lys Arg Thr His Ser Pro Ser Ser Lys Asp
1175 1180 1185
Glu Gln Ser Ile Gly Leu Lys Asp Ser Leu Leu Ala His Ser Ser
1190 1195 1200
Asp Pro Val Glu Met Arg Arg Leu Asn Tyr Gln Thr Pro Gly Met
1205 1210 1215
Arg Asp His Pro Pro Ile Pro Ile Thr Asp Leu Ala Asp Asn Ile
1220 1225 1230
Glu Arg Leu Lys Ala Asn Asp Gly Leu Lys Phe Ser Gln Glu Tyr
1235 1240 1245
Glu Ser Ile Asp Pro Gly Gln Gln Phe Thr Trp Glu Asn Ser Asn
1250 1255 1260
Leu Glu Val Asn Lys Pro Lys Asn Arg Tyr Ala Asn Val Ile Ala
1265 1270 1275
Tyr Asp His Ser Arg Val Ile Leu Thr Ser Ile Asp Gly Val Pro
1280 1285 1290
Gly Ser Asp Tyr Ile Asn Ala Asn Tyr Ile Asp Gly Tyr Arg Lys
1295 1300 1305
Gln Asn Ala Tyr Ile Ala Thr Gln Gly Pro Leu Pro Glu Thr Met
1310 1315 1320
Gly Asp Phe Trp Arg Met Val Trp Glu Gln Arg Thr Ala Thr Val
1325 1330 1335
Val Met Met Thr Arg Leu Glu Glu Lys Ser Arg Val Lys Cys Asp
1340 1345 1350
Gln Tyr Trp Pro Ala Arg Gly Thr Glu Thr Cys Gly Leu Ile Gln
1355 1360 1365
Val Thr Leu Leu Asp Thr Val Glu Leu Ala Thr Tyr Thr Val Arg
1370 1375 1380
Thr Phe Ala Leu His Lys Ser Gly Ser Ser Glu Lys Arg Glu Leu
1385 1390 1395
Arg Gln Phe Gln Phe Met Ala Trp Pro Asp His Gly Val Pro Glu
1400 1405 1410

```
Tyr Pro Thr Pro Ile Leu Ala Phe Leu Arg Arg Val Lys Ala Cys
               1415                1420                1425
Asn Pro Leu Asp Ala Gly Pro Met Val Val His Cys Ser Ala Gly
               1430                1435                1440
Val Gly Arg Thr Gly Cys Phe Ile Val Ile Asp Ala Met Leu Glu
               1445                1450                1455
Arg Met Lys His Glu Lys Thr Val Asp Ile Tyr Gly His Val Thr
               1460                1465                1470
Cys Met Arg Ser Gln Arg Asn Tyr Met Val Gln Thr Glu Asp Gln
               1475                1480                1485
Tyr Val Phe Ile His Glu Ala Leu Leu Glu Ala Ala Thr Cys Gly
               1490                1495                1500
His Thr Glu Val Pro Ala Arg Asn Leu Tyr Ala His Ile Gln Lys
               1505                1510                1515
Leu Gly Gln Val Pro Pro Gly Glu Ser Val Thr Ala Met Glu Leu
               1520                1525                1530
Glu Phe Lys Leu Leu Ala Ser Ser Lys Ala His Thr Ser Arg Phe
               1535                1540                1545
Ile Ser Ala Asn Leu Pro Cys Asn Lys Phe Lys Asn Arg Leu Val
               1550                1555                1560
Asn Ile Met Pro Tyr Glu Leu Thr Arg Val Cys Leu Gln Pro Ile
               1565                1570                1575
Arg Gly Val Glu Gly Ser Asp Tyr Ile Asn Ala Ser Phe Leu Asp
               1580                1585                1590
Gly Tyr Arg Gln Gln Lys Ala Tyr Ile Ala Thr Gln Gly Pro Leu
               1595                1600                1605
Ala Glu Ser Thr Glu Asp Phe Trp Arg Met Leu Trp Glu His Asn
               1610                1615                1620
Ser Thr Ile Ile Val Met Leu Thr Lys Leu Arg Glu Met Gly Arg
               1625                1630                1635
Glu Lys Cys His Gln Tyr Trp Pro Ala Glu Arg Ser Ala Arg Tyr
               1640                1645                1650
Gln Tyr Phe Val Val Asp Pro Met Ala Glu Tyr Asn Met Pro Gln
               1655                1660                1665
Tyr Ile Leu Arg Glu Phe Lys Val Thr Asp Ala Arg Asp Gly Gln
               1670                1675                1680
Ser Arg Thr Ile Arg Gln Phe Gln Phe Thr Asp Trp Pro Glu Gln
               1685                1690                1695
Gly Val Pro Lys Thr Gly Glu Gly Phe Ile Asp Phe Ile Gly Gln
               1700                1705                1710
Val His Lys Thr Lys Glu Gln Phe Gly Gln Asp Gly Pro Ile Thr
               1715                1720                1725
Val His Cys Ser Ala Gly Val Gly Arg Thr Gly Val Phe Ile Thr
               1730                1735                1740
Leu Ser Ile Val Leu Glu Arg Met Arg Tyr Glu Gly Val Val Asp
               1745                1750                1755
Met Phe Gln Thr Val Lys Thr Leu Arg Thr Gln Arg Pro Ala Met
               1760                1765                1770
Val Gln Thr Glu Asp Gln Tyr Gln Leu Cys Tyr Arg Ala Ala Leu
               1775                1780                1785
Glu Tyr Leu Gly Ser Phe Asp His Tyr Ala Thr
               1790                1795
```

<210> 17
<211> 438
<212> PRT

```
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6818489CD1

<400> 17
Met His Pro Leu Pro Gly Tyr Trp Ser Cys Tyr Cys Leu Leu Leu
  1               5                  10                  15
Leu Phe Ser Leu Gly Val Gln Gly Ser Leu Gly Ala Pro Ser Ala
                 20                  25                  30
Ala Pro Glu Gln Val His Leu Ser Tyr Pro Gly Glu Pro Gly Ser
                 35                  40                  45
Met Thr Val Thr Trp Thr Thr Trp Val Pro Thr Arg Ser Glu Val
                 50                  55                  60
Gln Phe Gly Leu Gln Pro Ser Gly Pro Leu Pro Leu Arg Ala Gln
                 65                  70                  75
Gly Thr Phe Val Pro Phe Val Asp Gly Gly Ile Leu Arg Arg Lys
                 80                  85                  90
Leu Tyr Ile His Arg Val Thr Leu Arg Lys Leu Leu Pro Gly Val
                 95                 100                 105
Gln Tyr Val Tyr Arg Cys Gly Ser Ala Gln Gly Trp Ser Arg Arg
                110                 115                 120
Phe Arg Phe Arg Ala Leu Lys Asn Gly Ala His Trp Ser Pro Arg
                125                 130                 135
Leu Ala Val Phe Gly Asp Leu Gly Ala Asp Asn Pro Lys Ala Val
                140                 145                 150
Pro Arg Leu Arg Arg Asp Thr Gln Gln Gly Met Tyr Asp Ala Val
                155                 160                 165
Leu His Val Gly Asp Phe Ala Tyr Asn Leu Asp Gln Asp Asn Ala
                170                 175                 180
Arg Val Gly Asp Arg Phe Met Arg Leu Ile Glu Pro Val Ala Ala
                185                 190                 195
Ser Leu Pro Tyr Met Thr Cys Pro Gly Asn His Glu Glu Arg Tyr
                200                 205                 210
Asn Phe Ser Asn Tyr Lys Ala Arg Phe Ser Met Pro Gly Asp Asn
                215                 220                 225
Glu Gly Leu Trp Tyr Ser Trp Asp Leu Gly Pro Ala His Ile Ile
                230                 235                 240
Ser Phe Ser Thr Glu Val Tyr Phe Phe Leu His Tyr Gly Arg His
                245                 250                 255
Leu Val Gln Arg Gln Phe Arg Trp Leu Glu Ser Asp Leu Gln Lys
                260                 265                 270
Ala Asn Lys Asn Arg Ala Ala Arg Pro Trp Ile Ile Thr Met Gly
                275                 280                 285
His Arg Pro Met Tyr Cys Ser Asn Ala Asp Leu Asp Asp Cys Thr
                290                 295                 300
Arg His Glu Ser Lys Val Arg Lys Gly Leu Gln Gly Lys Leu Tyr
                305                 310                 315
Gly Leu Glu Asp Leu Phe Tyr Lys Tyr Gly Val Asp Leu Gln Leu
                320                 325                 330
Trp Ala His Glu His Ser Tyr Glu Arg Leu Trp Pro Ile Tyr Asn
                335                 340                 345
Tyr Gln Val Phe Asn Gly Ser Arg Glu Met Pro Tyr Thr Asn Pro
                350                 355                 360
Arg Gly Pro Val His Ile Ile Thr Gly Ser Ala Gly Cys Glu Glu
```

365 370 375
Arg Leu Thr Pro Phe Ala Val Phe Pro Arg Pro Trp Ser Ala Val
380 385 390
Arg Val Lys Glu Tyr Gly Tyr Thr Arg Leu His Ile Leu Asn Gly
395 400 405
Thr His Ile His Ile Gln Gln Val Ser Asp Asp Gln Asp Gly Lys
410 415 420
Ile Val Asp Asp Val Trp Val Val Arg Pro Leu Phe Gly Arg Arg
425 430 435
Met Tyr Leu

<210> 18
<211> 635
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509415CD1

<400> 18
Met Gln Cys Leu Ala Ala Ala Leu Lys Asp Glu Thr Asn Met Ser
1 5 10 15
Gly Gly Gly Glu Gln Ala Asp Ile Leu Pro Ala Asn Tyr Val Val
20 25 30
Lys Asp Arg Trp Lys Val Leu Lys Lys Ile Gly Gly Gly Gly Phe
35 40 45
Gly Glu Ile Tyr Glu Ala Met Asp Leu Leu Thr Arg Glu Asn Val
50 55 60
Ala Leu Lys Val Glu Ser Ala Gln Gln Pro Lys Gln Val Leu Lys
65 70 75
Met Glu Val Ala Val Leu Lys Lys Leu Gln Gly Lys Asp His Val
80 85 90
Cys Arg Phe Ile Gly Cys Gly Arg Asn Glu Lys Phe Asn Tyr Val
95 100 105
Val Met Gln Leu Gln Gly Arg Asn Leu Ala Asp Leu Arg Arg Ser
110 115 120
Gln Pro Arg Gly Thr Phe Thr Leu Ser Thr Thr Leu Arg Leu Gly
125 130 135
Lys Gln Thr Leu Glu Ser Ile Glu Ala Ile His Ser Val Gly Phe
140 145 150
Leu His Arg Asp Ile Lys Pro Ser Asn Phe Ala Met Gly Arg Leu
155 160 165
Pro Ser Thr Tyr Arg Lys Cys Tyr Met Leu Asp Phe Gly Leu Ala
170 175 180
Arg Gln Tyr Thr Asn Thr Thr Gly Asp Val Arg Pro Pro Arg Asn
185 190 195
Val Ala Gly Phe Arg Gly Thr Val Arg Tyr Ala Ser Val Asn Ala
200 205 210
His Lys Asn Arg Glu Met Gly Arg His Asp Asp Leu Trp Ser Leu
215 220 225
Phe Tyr Met Leu Val Glu Phe Ala Val Gly Gln Leu Pro Trp Arg
230 235 240
Lys Ile Lys Asp Lys Glu Gln Val Gly Met Ile Lys Glu Lys Tyr
245 250 255

```
Glu His Arg Met Leu Leu Lys His Met Pro Ser Glu Phe His Leu
                260                 265                 270
Phe Leu Asp His Ile Ala Ser Leu Asp Tyr Phe Thr Lys Pro Asp
                275                 280                 285
Tyr Gln Leu Ile Met Ser Val Phe Glu Asn Ser Met Lys Glu Arg
                290                 295                 300
Gly Ile Ala Glu Asn Glu Ala Phe Asp Trp Glu Lys Ala Gly Thr
                305                 310                 315
Asp Ala Leu Leu Ser Thr Ser Thr Ser Thr Pro Pro Gln Gln Asn
                320                 325                 330
Thr Arg Gln Thr Ala Ala Met Phe Gly Val Val Asn Val Thr Pro
                335                 340                 345
Val Pro Gly Asp Leu Leu Arg Glu Asn Thr Glu Asp Val Leu Gln
                350                 355                 360
Gly Glu His Leu Ser Asp Gln Glu Asn Ala Pro Pro Ile Leu Pro
                365                 370                 375
Gly Arg Pro Ser Glu Gly Leu Gly His Ser Pro His Leu Val Pro
                380                 385                 390
His Pro Gly Gly Pro Glu Ala Glu Val Trp Glu Glu Thr Asp Val
                395                 400                 405
Asn Arg Asn Lys Leu Arg Ile Asn Ile Gly Lys Val Thr Ala Ala
                410                 415                 420
Arg Ala Lys Gly Val Gly Gly Leu Phe Ser His Pro Arg Phe Pro
                425                 430                 435
Ala Leu Cys Pro Cys Pro Val Pro Pro Lys His Pro Val Pro Gly
                440                 445                 450
His Leu Pro Ala Cys Pro Ala Ser Val Ser Arg Ser Leu Pro Ala
                455                 460                 465
Leu Ala Ser Leu Cys Leu Pro Ser Ser Ser Ser Ser Val Ser Phe
                470                 475                 480
Thr Leu Arg Arg Pro Ser Ala His Ser Arg Leu Ile Ser Pro Ser
                485                 490                 495
Ser Trp His Ser Pro Leu Leu Gln Ser Pro Cys Val Glu Glu Glu
                500                 505                 510
Gln Ser Arg Gly Met Gly Val Pro Ser Ser Pro Val Arg Ala Pro
                515                 520                 525
Pro Asp Ser Pro Thr Thr Pro Val Arg Ser Leu Arg Tyr Arg Arg
                530                 535                 540
Val Asn Ser Pro Glu Ser Glu Arg Leu Ser Thr Ala Asp Gly Arg
                545                 550                 555
Val Glu Leu Pro Glu Arg Arg Trp Val Trp Gly Gln Gly His Gly
                560                 565                 570
Trp Gly Pro Arg Pro Ser Pro Pro Ser Arg Gly Trp Ser Gly Gly
                575                 580                 585
Lys Val Arg Cys Val Ala Glu Val Gly Arg Pro Trp Glu Val Leu
                590                 595                 600
Arg Gly Leu Tyr Leu Gly Leu Gly Ser Asp Ser Val Gly Ala Arg
                605                 610                 615
Asp Arg Ala Trp Glu Asn Gln Trp Gly Ile Gln Arg Gly Pro Gly
                620                 625                 630
Ser Cys Gln Glu Thr
                635

<210> 19
<211> 444
<212> PRT
```

```
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506916CD1

<400> 19
Met Arg Ser Gly Ala Glu Arg Arg Gly Ser Ser Ala Ala Ala Ser
1               5                   10                  15
Pro Gly Ser Pro Pro Pro Gly Arg Ala Arg Pro Ala Gly Ser Asp
                20                  25                  30
Ala Pro Ser Ala Leu Pro Pro Pro Ala Ala Gly Gln Pro Arg Ala
                35                  40                  45
Arg Asp Ser Gly Asp Val Arg Ser Gln Pro Arg Pro Leu Phe Gln
                50                  55                  60
Trp Ser Lys Trp Lys Lys Arg Met Gly Ser Ser Met Ser Ala Ala
                65                  70                  75
Thr Ala Arg Arg Pro Val Phe Asp Asp Lys Glu Asp Val Asn Phe
                80                  85                  90
Asp His Phe Gln Ile Leu Arg Ala Ile Gly Lys Gly Ser Phe Gly
                95                  100                 105
Lys Val Cys Ile Val Gln Lys Arg Asp Thr Glu Lys Met Tyr Ala
                110                 115                 120
Met Lys Tyr Met Asn Lys Gln Gln Cys Ile Glu Arg Asp Glu Val
                125                 130                 135
Arg Asn Val Phe Arg Glu Leu Glu Ile Leu Gln Glu Ile Glu His
                140                 145                 150
Val Phe Leu Val Asn Leu Trp Tyr Ser Phe Gln Asp Glu Glu Asp
                155                 160                 165
Met Phe Met Val Val Asp Leu Leu Leu Gly Gly Asp Leu Arg Tyr
                170                 175                 180
His Leu Gln Gln Asn Val Gln Phe Ser Glu Asp Thr Val Arg Leu
                185                 190                 195
Tyr Ile Cys Glu Met Ala Leu Ala Leu Asp Tyr Leu Arg Gly Gln
                200                 205                 210
His Ile Ile His Arg Asp Val Lys Pro Asp Asn Ile Leu Leu Asp
                215                 220                 225
Glu Arg Gly His Ala His Leu Thr Asp Phe Asn Ile Ala Thr Ile
                230                 235                 240
Ile Lys Asp Gly Glu Arg Ala Thr Ala Leu Ala Gly Thr Lys Pro
                245                 250                 255
Tyr Met Ala Pro Glu Ile Phe His Ser Phe Val Asn Gly Gly Thr
                260                 265                 270
Gly Tyr Ser Phe Glu Val Asp Trp Trp Ser Val Gly Val Met Ala
                275                 280                 285
Tyr Glu Leu Leu Arg Gly Trp Arg Pro Tyr Asp Ile His Ser Ser
                290                 295                 300
Asn Ala Val Glu Ser Leu Val Gln Leu Phe Ser Thr Val Ser Val
                305                 310                 315
Gln Tyr Val Pro Thr Trp Ser Lys Glu Met Val Ala Leu Leu Arg
                320                 325                 330
Lys Lys Gly Arg Leu His Cys Asp Pro Thr Phe Glu Leu Glu Glu
                335                 340                 345
Met Ile Leu Glu Ser Arg Pro Leu His Lys Lys Lys Lys Arg Leu
                350                 355                 360
Ala Lys Asn Lys Ser Arg Asp Asn Ser Arg Asp Ser Ser Gln Ser
```

```
                365                 370                 375
Glu Asn Asp Tyr Leu Gln Asp Cys Leu Asp Ala Ile Gln Gln Asp
                380                 385                 390
Phe Val Ile Phe Asn Arg Glu Lys Leu Lys Arg Ser Gln Asp Leu
                395                 400                 405
Pro Arg Glu Pro Leu Pro Ala Pro Glu Ser Arg Asp Ala Ala Glu
                410                 415                 420
Pro Val Glu Asp Glu Ala Glu Arg Ser Ala Leu Pro Met Cys Gly
                425                 430                 435
Pro Ile Cys Pro Ser Ala Gly Ser Gly
                440

<210> 20
<211> 230
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507104CD1

<400> 20
Met Pro Ser Ser Ser Trp Ala Ile Ile Val Arg Asn Gly Glu Ile
  1             5                   10                  15
Ile Pro Met Ser Arg Glu Phe Thr Pro Glu Thr Glu Arg Gln Arg
                20                  25                  30
Leu Gln Leu Leu Gly Phe Leu Lys Pro Glu Leu Leu Gly Ser Glu
                35                  40                  45
Phe Thr His Leu Glu Phe Pro Arg Arg Val Leu Pro Lys Glu Leu
                50                  55                  60
Gly Gln Arg Met Leu Tyr Arg Asp Gln Asn Met Thr Gly Trp Ala
                65                  70                  75
Tyr Lys Lys Ile Glu Leu Glu Asp Leu Arg Phe Pro Leu Val Cys
                80                  85                  90
Gly Glu Gly Lys Lys Ala Arg Val Met Ala Thr Ile Gly Val Thr
                95                  100                 105
Arg Gly Leu Gly Asp His Ser Leu Lys Val Cys Ser Ser Thr Leu
                110                 115                 120
Pro Ile Lys Pro Phe Leu Ser Cys Phe Pro Glu Val Arg Val Tyr
                125                 130                 135
Asp Leu Thr Gln Tyr Glu His Cys Pro Asp Asp Val Leu Val Leu
                140                 145                 150
Gly Thr Asp Gly Leu Trp Asp Val Thr Thr Asp Cys Glu Val Ala
                155                 160                 165
Ala Thr Val Asp Arg Val Leu Ser Ala Tyr Glu Pro Asn Asp His
                170                 175                 180
Ser Arg Tyr Thr Ala Leu Ala Gln Ala Leu Val Leu Gly Ala Arg
                185                 190                 195
Gly Thr Pro Arg Asp Arg Gly Trp Arg Leu Pro Asn Asn Lys Leu
                200                 205                 210
Gly Ser Gly Asp Asp Ile Ser Val Phe Val Ile Pro Leu Gly Gly
                215                 220                 225
Pro Gly Ser Tyr Ser
                230

<210> 21
```

```
<211> 320
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507105CD1

<400> 21
Met Ala Ser Arg Leu Leu His Arg His Ile Arg Glu Gln Leu Lys
 1              5                   10                  15
Asp Leu Val Glu Ile Leu Gln Asp Pro Ser Pro Pro Pro Leu Cys
                20                  25                  30
Leu Pro Thr Thr Pro Gly Thr Pro Asp Ser Ser Asp Pro Ser His
                35                  40                  45
Leu Leu Gly Pro Gln Ser Cys Trp Ser Ser Gln Lys Glu Asp Glu
                50                  55                  60
Gln Met Ala Arg Glu Arg Arg Gly His Gln Val Glu Gly Gly Cys
                65                  70                  75
Cys Ala Leu Val Val Ile Tyr Leu Leu Gly Lys Val Tyr Val Ala
                80                  85                  90
Asn Ala Gly Asp Ser Arg Ala Ile Ile Val Arg Asn Gly Glu Ile
                95                  100                 105
Ile Pro Met Ser Arg Glu Phe Thr Pro Glu Thr Glu Arg Gln Arg
                110                 115                 120
Leu Gln Leu Leu Gly Phe Leu Lys Pro Glu Leu Leu Gly Ser Glu
                125                 130                 135
Phe Thr His Leu Glu Phe Pro Arg Arg Val Leu Pro Lys Glu Leu
                140                 145                 150
Gly Gln Arg Met Leu Tyr Arg Asp Gln Asn Met Thr Gly Trp Ala
                155                 160                 165
Tyr Lys Lys Ile Glu Leu Glu Asp Leu Arg Phe Pro Leu Val Cys
                170                 175                 180
Gly Glu Gly Lys Lys Ala Arg Val Met Ala Thr Ile Gly Val Thr
                185                 190                 195
Arg Gly Leu Gly Asp His Ser Leu Lys Val Cys Ser Ser Thr Leu
                200                 205                 210
Pro Ile Lys Pro Phe Leu Ser Cys Phe Pro Glu Val Arg Val Tyr
                215                 220                 225
Asp Leu Thr Gln Tyr Glu His Cys Pro Asp Asp Val Leu Val Leu
                230                 235                 240
Gly Thr Asp Gly Leu Trp Asp Val Thr Thr Asp Cys Glu Val Ala
                245                 250                 255
Ala Thr Val Asp Arg Val Leu Ser Ala Tyr Glu Pro Asn Asp His
                260                 265                 270
Ser Arg Tyr Thr Ala Leu Ala Gln Ala Leu Val Leu Gly Ala Arg
                275                 280                 285
Gly Thr Pro Arg Asp Arg Gly Trp Arg Leu Pro Asn Asn Lys Leu
                290                 295                 300
Gly Ser Gly Asp Asp Ile Ser Val Phe Val Ile Pro Leu Gly Gly
                305                 310                 315
Pro Gly Ser Tyr Ser
                320

<210> 22
<211> 184
```

```
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507107CD1

<400> 22
Met Ala Ser Arg Leu Leu His Arg His Ile Arg Glu Gln Leu Lys
 1               5                  10                  15
Asp Leu Val Glu Ile Leu Gln Asp Pro Ser Pro Pro Pro Leu Cys
                20                  25                  30
Leu Pro Thr Thr Pro Gly Thr Pro Asp Ser Ser Asp Pro Ser His
                35                  40                  45
Leu Leu Gly Pro Gln Ser Cys Trp Ser Ser Gln Lys Glu Val Ser
                50                  55                  60
His Glu Ser Leu Val Val Gly Ala Ile Glu Asn Ala Phe Gln Leu
                65                  70                  75
Met Asp Glu Gln Met Ala Arg Glu Arg Arg Gly His Gln Val Glu
                80                  85                  90
Gly Gly Cys Cys Ala Leu Val Val Ile Tyr Leu Leu Gly Lys Val
                95                  100                 105
Tyr Val Ala Asn Ala Gly Asp Ser Arg Ala Ile Ile Val Arg Asn
                110                 115                 120
Gly Glu Ile Ile Pro Met Ser Arg Glu Phe Thr Pro Glu Thr Glu
                125                 130                 135
Arg Gln Arg Leu Gln Leu Leu Gly Phe Leu Lys Pro Glu Leu Leu
                140                 145                 150
Gly Ser Glu Phe Thr His Leu Glu Phe Pro Arg Arg Val Leu Pro
                155                 160                 165
Lys Glu Leu Gly Gln Arg Met Leu Tyr Arg Asp Gln Asn Met Thr
                170                 175                 180
Gly Trp Leu Gly


<210> 23
<211> 303
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507109CD1

<400> 23
Met Ala Ser Arg Leu Leu His Arg His Ile Arg Glu Gln Leu Lys
 1               5                  10                  15
Asp Leu Val Glu Ile Leu Gln Asp Pro Ser Pro Pro Pro Leu Cys
                20                  25                  30
Leu Pro Thr Thr Pro Gly Thr Pro Asp Ser Ser Asp Pro Ser His
                35                  40                  45
Leu Leu Gly Pro Gln Ser Cys Trp Ser Ser Gln Lys Glu Val Ser
                50                  55                  60
His Glu Ser Leu Val Val Gly Ala Ile Glu Asn Ala Phe Gln Leu
                65                  70                  75
Met Asp Glu Gln Met Ala Arg Glu Arg Arg Gly His Gln Val Glu
```

```
                 80                  85                  90
Gly Gly Cys Cys Ala Leu Val Val Ile Tyr Leu Leu Gly Lys Val
                 95                 100                 105
Tyr Val Ala Asn Ala Gly Asp Ser Arg Ala Ile Ile Val Arg Asn
                110                 115                 120
Gly Glu Ile Ile Pro Met Ser Arg Glu Phe Thr Pro Glu Thr Glu
                125                 130                 135
Arg Gln Arg Leu Gln Leu Leu Gly Phe Leu Lys Pro Glu Leu Leu
                140                 145                 150
Gly Ser Glu Phe Thr His Leu Glu Phe Pro Arg Arg Val Leu Pro
                155                 160                 165
Lys Glu Leu Gly Gln Arg Met Leu Tyr Arg Asp Gln Asn Met Thr
                170                 175                 180
Gly Trp Ala Tyr Lys Lys Ile Glu Leu Glu Asp Leu Arg Phe Pro
                185                 190                 195
Leu Val Cys Gly Glu Gly Lys Lys Ala Arg Val Met Ala Thr Ile
                200                 205                 210
Gly Val Thr Arg Gly Leu Gly Asp His Ser Leu Lys Val Cys Ser
                215                 220                 225
Ser Thr Leu Pro Ile Lys Pro Phe Leu Ser Cys Phe Pro Glu Val
                230                 235                 240
Arg Val Tyr Asp Leu Thr Gln Tyr Glu His Cys Pro Asp Asp Val
                245                 250                 255
Leu Val Leu Gly Thr Asp Gly Leu Trp Asp Val Thr Thr Asp Cys
                260                 265                 270
Glu Val Tyr Ser Ser Gly Pro Ser Ser Gly Pro Gly Gly Pro Gly
                275                 280                 285
Tyr Pro Pro Arg Pro Trp Leu Ala Ser Pro Gln Gln Gln Ala Gly
                290                 295                 300
Phe Arg Gly


<210> 24
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1833937CD1

<400> 24
Met Ala Ala Gly Arg Leu Phe Leu Ser Arg Leu Arg Ala Pro Phe
 1                5                  10                  15
Ser Ser Met Ala Lys Ser Pro Leu Glu Gly Val Ser Ser Ser Arg
                 20                  25                  30
Gly Leu His Ala Gly Arg Gly Pro Arg Arg Leu Ser Ile Glu Gly
                 35                  40                  45
Asn Ile Gly Ser Thr Leu Arg Leu
                 50

<210> 25
<211> 495
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 7502036CD1

<400> 25
Met Ser Thr Ser Glu Thr Ser Lys His Arg Val Thr Gly Gly Glu
1               5                   10                  15
Leu Phe Glu Asp Ile Val Ala Arg Glu Tyr Tyr Ser Glu Ala Asp
                20                  25                  30
Ala Ser His Cys Ile His Gln Ile Leu Glu Ser Val Asn His Ile
                35                  40                  45
His Gln His Asp Ile Val His Arg Asp Leu Lys Pro Glu Asn Leu
                50                  55                  60
Leu Leu Ala Ser Lys Cys Lys Gly Ala Ala Val Lys Leu Ala Asp
                65                  70                  75
Phe Gly Leu Ala Ile Glu Val Gln Gly Glu Gln Gln Ala Trp Phe
                80                  85                  90
Gly Phe Ala Gly Thr Pro Gly Tyr Leu Ser Pro Glu Val Leu Arg
                95                  100                 105
Lys Asp Pro Tyr Gly Lys Pro Val Asp Ile Trp Ala Cys Gly Val
                110                 115                 120
Ile Leu Tyr Ile Leu Leu Val Gly Tyr Pro Pro Phe Trp Asp Glu
                125                 130                 135
Asp Gln His Lys Leu Tyr Gln Gln Ile Lys Ala Gly Ala Tyr Asp
                140                 145                 150
Phe Pro Ser Pro Glu Trp Asp Thr Val Thr Pro Glu Ala Lys Asn
                155                 160                 165
Leu Ile Asn Gln Met Leu Thr Ile Asn Pro Ala Lys Arg Ile Thr
                170                 175                 180
Ala Asp Gln Ala Leu Lys His Pro Trp Val Cys Gln Arg Ser Thr
                185                 190                 195
Val Ala Ser Met Met His Arg Gln Glu Thr Val Glu Cys Leu Arg
                200                 205                 210
Lys Phe Asn Ala Arg Arg Lys Leu Lys Gly Ala Ile Leu Thr Thr
                215                 220                 225
Met Leu Val Ser Arg Asn Phe Ser Val Gly Arg Gln Ser Ser Ala
                230                 235                 240
Pro Ala Ser Pro Ala Ala Ser Ala Ala Gly Leu Ala Gly Gln Ala
                245                 250                 255
Ala Lys Ser Leu Leu Asn Lys Lys Ser Asp Gly Gly Val Lys Pro
                260                 265                 270
Gln Ser Asn Asn Lys Asn Ser Leu Val Ser Pro Ala Gln Glu Pro
                275                 280                 285
Ala Pro Leu Gln Thr Ala Met Glu Pro Gln Thr Thr Val Val His
                290                 295                 300
Asn Ala Thr Asp Gly Ile Lys Gly Ser Thr Glu Ser Cys Asn Thr
                305                 310                 315
Thr Thr Glu Asp Glu Asp Leu Lys Ala Ala Pro Leu Arg Thr Gly
                320                 325                 330
Asn Gly Ser Ser Val Pro Glu Gly Arg Ser Ser Arg Asp Arg Thr
                335                 340                 345
Ala Pro Ser Ala Gly Met Gln Pro Gln Pro Ser Leu Cys Ser Ser
                350                 355                 360
Ala Met Arg Lys Gln Glu Ile Ile Lys Ile Thr Glu Gln Leu Ile
                365                 370                 375
Glu Ala Ile Asn Asn Gly Asp Phe Glu Ala Tyr Thr Lys Ile Cys
```

```
                380                 385                 390
Asp Pro Gly Leu Thr Ser Phe Glu Pro Glu Ala Leu Gly Asn Leu
                395                 400                 405
Val Glu Gly Met Asp Phe His Lys Phe Tyr Phe Glu Asn Leu Leu
                410                 415                 420
Ser Lys Asn Ser Lys Pro Ile His Thr Thr Ile Leu Asn Pro His
                425                 430                 435
Val His Val Ile Gly Glu Asp Ala Ala Cys Ile Ala Tyr Ile Arg
                440                 445                 450
Leu Thr Gln Tyr Ile Asp Gly Gln Gly Arg Pro Arg Thr Ser Gln
                455                 460                 465
Ser Glu Glu Thr Arg Val Trp His Arg Arg Asp Gly Lys Trp Leu
                470                 475                 480
Asn Val His Tyr His Cys Ser Gly Ala Pro Ala Ala Pro Leu Gln
                485                 490                 495


<210> 26
<211> 221
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503248CD1

<400> 26
Met Ala Pro Val Tyr Glu Gly Met Ala Ser His Val Gln Val Phe
 1               5                   10                  15
Ser Pro His Thr Leu Gln Ser Ser Ala Phe Cys Ser Val Lys Lys
                 20                  25                  30
Leu Lys Ile Glu Pro Ser Ser Asn Trp Asp Met Thr Gly Tyr Gly
                 35                  40                  45
Ser His Ser Lys Val Tyr Ser Gln Ser Lys Asn Ile Pro Leu Ser
                 50                  55                  60
Gln Pro Ala Thr Thr Thr Val Ser Thr Ser Leu Pro Val Pro Asn
                 65                  70                  75
Pro Ser Leu Pro Tyr Glu Gln Thr Ile Val Phe Pro Gly Ser Thr
                 80                  85                  90
Gly His Ile Val Val Thr Ser Ala Ser Ser Thr Ser Val Thr Gly
                 95                 100                 105
Gln Val Leu Gly Gly Pro His Asn Leu Met Arg Arg Ser Thr Val
                110                 115                 120
Ser Leu Leu Asp Thr Tyr Gln Lys Cys Gly Leu Lys Arg Lys Ser
                125                 130                 135
Glu Glu Ile Glu Asn Thr Ser Ser Val Gln Ile Ile Glu Glu His
                140                 145                 150
Pro Pro Met Ile Gln Asn Asn Ala Ser Gly Ala Thr Val Ala Thr
                155                 160                 165
Ala Thr Thr Ser Thr Ala Thr Ser Lys Asn Ser Gly Ser Pro Thr
                170                 175                 180
Ile His Pro Ser Gln Tyr Pro Ala Gln Phe Ala His Gln Thr Tyr
                185                 190                 195
Ile Ser Ala Ser Pro Ala Ser Thr Val Tyr Thr Gly Tyr Pro Leu
                200                 205                 210
Ser Pro Ala Lys Val Asn Gln Tyr Pro Tyr Ile
```

```
                215                 220

<210> 27
<211> 307
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503968CD1

<400> 27
Met Leu Gly Pro Gly Ser Asn Arg Arg Arg Pro Thr Gln Gly Glu
1               5                   10                  15
Arg Gly Pro Gly Ser Pro Gly Glu Pro Met Glu Lys Tyr Gln Val
                20                  25                  30
Leu Tyr Gln Leu Asn Pro Gly Ala Leu Gly Val Asn Leu Val Val
                35                  40                  45
Glu Glu Met Glu Thr Lys Val Lys His Val Ile Lys Gln Val Glu
                50                  55                  60
Cys Met Asp Asp His Tyr Ala Ser Gln Ala Leu Glu Glu Leu Met
                65                  70                  75
Pro Leu Leu Lys Leu Arg His Ala His Ile Ser Val Tyr Gln Glu
                80                  85                  90
Leu Phe Ile Thr Trp Asn Gly Glu Ile Ser Ser Leu Tyr Leu Cys
                95                  100                 105
Leu Val Met Glu Phe Asn Glu Leu Ser Phe Gln Glu Val Ile Glu
                110                 115                 120
Asp Lys Arg Lys Ala Lys Lys Ile Ile Asp Ser Glu Trp Met Gln
                125                 130                 135
Asn Val Leu Gly Gln Val Leu Asp Ala Leu Glu Tyr Leu His His
                140                 145                 150
Leu Asp Ile Ile His Arg Asn Leu Lys Pro Ser Asn Ile Ile Leu
                155                 160                 165
Ile Ser Ser Asp His Cys Lys Leu Gln Asp Leu Ser Ser Asn Val
                170                 175                 180
Leu Met Thr Asp Lys Ala Lys Trp Asn Ile Arg Ala Glu Glu Asp
                185                 190                 195
Pro Phe Arg Lys Ser Trp Met Ala Pro Glu Ala Leu Asn Phe Ser
                200                 205                 210
Phe Ser Gln Lys Ser Asp Ile Trp Ser Leu Gly Cys Ile Ile Leu
                215                 220                 225
Asp Met Thr Ser Cys Ser Phe Met Asp Gly Thr Glu Ala Met His
                230                 235                 240
Leu Arg Lys Ser Leu Arg Gln Ser Pro Gly Ser Leu Lys Ala Val
                245                 250                 255
Leu Lys Thr Met Glu Glu Lys Gln Ile Pro Asp Val Glu Thr Phe
                260                 265                 270
Arg Asn Leu Leu Pro Leu Met Leu Gln Ile Asp Pro Ser Asp Arg
                275                 280                 285
Ile Thr Ile Lys Ser Ala Val Ala Pro Gly Ala Gly Gly Gly Gly
                290                 295                 300
Gly His Asp His Gly Ala Thr
                305

<210> 28
```

```
<211> 142
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505931CD1

<400> 28
Met Ala Ala Ser Lys Lys Ala Val Leu Gly Pro Leu Val Gly Ala
  1               5                  10                  15
Val Asp Gln Gly Thr Ser Ser Thr Arg Phe Leu Val Phe Asn Ser
                 20                  25                  30
Lys Thr Ala Glu Leu Leu Ser His His Gln Val Glu Ile Lys Gln
                 35                  40                  45
Glu Phe Pro Arg Glu Gly Trp Val Glu Gln Asp Pro Lys Glu Ile
                 50                  55                  60
Leu His Ser Val Tyr Glu Cys Ile Glu Lys Thr Cys Glu Lys Leu
                 65                  70                  75
Gly Gln Leu Asn Ile Asp Ile Ser Asn Ile Lys Ala Ile Gly Val
                 80                  85                  90
Ser Asn Gln Arg Glu Thr Thr Val Val Trp Asp Lys Ile Thr Gly
                 95                 100                 105
Glu Pro Leu Tyr Asn Ala Val Glu Ser Glu Ile Arg Tyr Ser Thr
                110                 115                 120
Trp Lys Lys Ala Val Met Lys Ser Met Gly Trp Val Thr Thr Gln
                125                 130                 135
Ser Pro Glu Ser Gly Ile Pro
                140

<210> 29
<211> 388
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506912CD1

<400> 29
Met Ala Lys Ser Lys Gln Glu Gly Leu His Gln Met Gln Ser Leu
  1               5                  10                  15
Asp Leu Gly Leu Leu Ser Leu Gln Asn Ser Phe Pro Trp Phe Gly
                 20                  25                  30
Met Asp Ile Gly Gly Thr Leu Val Lys Leu Val Tyr Phe Glu Pro
                 35                  40                  45
Lys Asp Ile Thr Ala Glu Glu Glu Gln Glu Glu Val Glu Asn Leu
                 50                  55                  60
Lys Ser Ile Arg Lys Tyr Leu Thr Ser Asn Thr Ala Tyr Gly Lys
                 65                  70                  75
Thr Gly Ile Arg Asp Val His Leu Glu Leu Lys Asn Leu Thr Met
                 80                  85                  90
Cys Gly Arg Lys Gly Asn Leu His Phe Ile Arg Phe Pro Ser Cys
                 95                 100                 105
Ala Met His Arg Phe Ile Gln Met Gly Ser Glu Lys Asn Phe Ser
                110                 115                 120
```

```
Ser Leu His Thr Thr Leu Cys Ala Thr Gly Gly Gly Ala Phe Lys
                125                 130                 135
Phe Glu Glu Asp Phe Arg Met Ile Ala Asp Leu Gln Leu His Lys
                140                 145                 150
Leu Asp Glu Leu Asp Cys Leu Ile Gln Gly Leu Leu Tyr Val Asp
                155                 160                 165
Ser Val Gly Phe Asn Gly Lys Pro Glu Cys Tyr Tyr Phe Glu Asn
                170                 175                 180
Pro Thr Asn Pro Glu Leu Cys Gln Lys Lys Pro Tyr Cys Leu Asp
                185                 190                 195
Asn Pro Tyr Pro Met Leu Leu Val Asn Met Gly Ser Gly Val Ser
                200                 205                 210
Ile Leu Ala Val Tyr Ser Lys Asp Asn Tyr Lys Arg Val Thr Gly
                215                 220                 225
Thr Ser Leu Gly Gly Gly Thr Phe Leu Gly Leu Cys Cys Leu Leu
                230                 235                 240
Thr Gly Cys Glu Thr Phe Glu Glu Ala Leu Glu Met Ala Ala Lys
                245                 250                 255
Gly Asp Ser Thr Asn Val Asp Lys Leu Val Lys Asp Ile Tyr Gly
                260                 265                 270
Gly Asp Tyr Glu Arg Phe Gly Leu Gln Gly Ser Ala Val Ala Ser
                275                 280                 285
Ser Phe Gly Asn Met Met Ser Lys Glu Lys Arg Asp Ser Ile Ser
                290                 295                 300
Lys Glu Asp Leu Ala Arg Ala Thr Leu Val Thr Ile Thr Asn Asn
                305                 310                 315
Ile Gly Ser Ile Ala Arg Met Cys Ala Leu Asn Glu Asn Ile Asp
                320                 325                 330
Arg Val Val Phe Val Gly Asn Phe Leu Arg Ile Asn Met Val Ser
                335                 340                 345
Met Lys Leu Leu Ala Tyr Ala Met Asp Phe Trp Ser Lys Gly Gln
                350                 355                 360
Leu Lys Ala Leu Phe Leu Glu His Glu Gly Tyr Phe Gly Ala Val
                365                 370                 375
Gly Ala Leu Leu Glu Leu Phe Lys Met Thr Asp Asp Lys
                380                 385

<210> 30
<211> 314
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506913CD1

<400> 30
Met Lys Leu Ile Asn Gly Lys Lys Gln Thr Phe Pro Trp Phe Gly
  1               5                  10                  15
Met Asp Ile Gly Gly Thr Leu Val Lys Leu Val Tyr Phe Glu Pro
                 20                  25                  30
Lys Asp Ile Thr Ala Glu Glu Glu Gln Glu Glu Val Glu Asn Leu
                 35                  40                  45
Lys Ser Ile Arg Lys Tyr Leu Thr Ser Asn Thr Ala Tyr Gly Lys
                 50                  55                  60
Thr Gly Ile Arg Asp Val His Leu Glu Leu Lys Asn Leu Thr Met
```

```
                65                  70                  75
Cys Gly Arg Lys Gly Asn Leu His Phe Ile Arg Phe Pro Ser Cys
                80                  85                  90
Ala Met His Arg Phe Ile Gln Met Gly Ser Glu Lys Asn Phe Ser
                95                  100                 105
Ser Leu His Thr Thr Leu Cys Ala Thr Gly Gly Gly Ala Phe Lys
                110                 115                 120
Phe Glu Glu Asp Phe Arg Met Ile Ala Asp Leu Gln Leu His Lys
                125                 130                 135
Leu Asp Glu Leu Asp Cys Leu Ile Gln Gly Leu Leu Tyr Val Asp
                140                 145                 150
Ser Val Gly Phe Asn Gly Lys Pro Glu Cys Tyr Tyr Phe Glu Asn
                155                 160                 165
Pro Thr Asn Pro Glu Leu Cys Gln Lys Lys Pro Tyr Cys Leu Asp
                170                 175                 180
Asn Pro Tyr Pro Met Leu Leu Val Asn Met Gly Ser Gly Val Ser
                185                 190                 195
Ile Leu Ala Val Tyr Ser Lys Asp Asn Tyr Lys Arg Val Thr Gly
                200                 205                 210
Thr Ser Phe Gly Asn Met Met Ser Lys Glu Lys Arg Asp Ser Ile
                215                 220                 225
Ser Lys Glu Asp Leu Ala Arg Ala Thr Leu Val Thr Ile Thr Asn
                230                 235                 240
Asn Ile Gly Ser Ile Ala Arg Met Cys Ala Leu Asn Glu Asn Ile
                245                 250                 255
Asp Arg Val Val Phe Val Gly Asn Phe Leu Arg Ile Asn Met Val
                260                 265                 270
Ser Met Lys Leu Leu Ala Tyr Ala Met Asp Phe Trp Ser Lys Gly
                275                 280                 285
Gln Leu Lys Ala Leu Phe Leu Glu His Glu Gly Tyr Phe Gly Ala
                290                 295                 300
Val Gly Ala Leu Leu Glu Leu Phe Lys Met Thr Asp Asp Lys
                305                 310
```

<210> 31
<211> 166
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507029CD1

<400> 31

```
Met Tyr Ser Leu Asn Gln Glu Ile Lys Ala Phe Ser Arg Asn Asn
  1             5                   10                  15
Leu Arg Lys Gln Cys Thr Arg Val Thr Thr Leu Thr Gly Lys Lys
                20                  25                  30
Ile Ile Glu Thr Trp Lys Asp Ala Arg Ile His Val Val Glu Glu
                35                  40                  45
Val Glu Pro Ser Ser Gly Gly Gly Cys Gly Tyr Val Gln Asp Leu
                50                  55                  60
Ser Ser Asp Leu Gln Val Gly Val Ile Lys Pro Trp Leu Leu Leu
                65                  70                  75
Gly Ser Gln Asp Ala Ala His Asp Leu Asp Thr Leu Lys Lys Asn
                80                  85                  90
```

```
Lys Asp Gly Val Val Leu Val His Cys Asn Ala Gly Val Ser Arg
                 95                  100                 105
Ala Ala Ala Ile Val Ile Gly Phe Leu Met Asn Ser Glu Gln Thr
                110                  115                 120
Ser Phe Thr Ser Ala Phe Ser Leu Val Lys Asn Ala Arg Pro Ser
                125                  130                 135
Ile Cys Pro Asn Ser Gly Phe Met Glu Gln Leu Arg Thr Tyr Gln
                140                  145                 150
Glu Gly Lys Glu Ser Asn Lys Cys Asp Arg Ile Gln Glu Asn Ser
                155                  160                 165
Ser


<210> 32
<211> 173
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507063CD1

<400> 32
Met Pro Cys Lys Ser Ala Glu Trp Leu Gln Glu Glu Leu Glu Ala
 1                5                   10                  15
Arg Gly Gly Ala Ser Leu Leu Leu Leu Asp Cys Arg Pro His Glu
                 20                   25                  30
Leu Phe Glu Ser Ser His Ile Glu Thr Ala Ile Asn Leu Ala Ile
                 35                   40                  45
Pro Gly Leu Met Leu Arg Arg Leu Arg Lys Gly Asn Leu Pro Ile
                 50                   55                  60
Arg Ser Ile Ile Pro Asn His Ala Asp Lys Glu Arg Phe Ala Thr
                 65                   70                  75
Arg Cys Lys Ala Ala Thr Val Leu Leu Tyr Asp Glu Ala Thr Ala
                 80                   85                  90
Glu Trp Gln Pro Glu Pro Gly Ala Pro Ala Ser Val Leu Gly Leu
                 95                  100                 105
Leu Leu Gln Lys Leu Arg Asp Asp Gly Cys Gln Ala Tyr Tyr Leu
                110                  115                 120
Gln Asp Pro Ala Leu Pro Leu Pro Arg Leu Arg Gln Gly Leu His
                125                  130                 135
Gln Pro Gly Arg Ala Arg Gln Val Trp His Gln Val Tyr Pro Gln
                140                  145                 150
Cys His Thr Gln Pro Thr Gln Arg Leu Arg Ala Arg Arg Arg Val
                155                  160                 165
His Leu Gln Ala Asp Pro His Leu
                170

<210> 33
<211> 486
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504755CD1
```

```
<400> 33
Met Gly Cys Trp Gly Gln Leu Leu Val Trp Phe Gly Ala Ala Gly
 1               5                  10                  15
Ala Ile Leu Cys Ser Ser Pro Gly Ser Gln Glu Thr Phe Leu Arg
                20                  25                  30
Ser Ser Pro Leu Pro Leu Ala Ser Pro Ser Pro Gln Asp Pro Lys
                35                  40                  45
Val Ser Ala Pro Pro Ser Ile Leu Glu Pro Ala Ser Pro Leu Asn
                50                  55                  60
Ser Pro Gly Thr Glu Gly Ser Trp Leu Phe Ser Thr Cys Gly Ala
                65                  70                  75
Ser Gly Arg His Gly Pro Thr Gln Thr Gln Cys Asp Gly Ala Tyr
                80                  85                  90
Ala Gly Thr Ser Val Val Val Thr Val Gly Ala Ala Gly Gln Leu
                95                  100                 105
Arg Gly Val Gln Leu Trp Arg Val Pro Gly Pro Gly Gln Tyr Leu
                110                 115                 120
Ile Ser Ala Tyr Gly Ala Ala Gly Gly Lys Gly Ala Lys Asn His
                125                 130                 135
Leu Ser Arg Ala His Gly Val Phe Val Ser Ala Ile Phe Ser Leu
                140                 145                 150
Gly Leu Gly Glu Ser Leu Tyr Ile Leu Val Gly Gln Gln Gly Glu
                155                 160                 165
Asp Ala Cys Pro Gly Gly Ser Pro Glu Ser Gln Leu Val Cys Leu
                170                 175                 180
Gly Glu Ser Arg Ala Val Glu Glu His Ala Ala Met Asp Gly Ser
                185                 190                 195
Glu Gly Val Pro Gly Ser Arg Arg Trp Ala Gly Gly Gly Gly Gly
                200                 205                 210
Gly Gly Gly Ala Thr Tyr Val Phe Arg Val Arg Ala Gly Glu Leu
                215                 220                 225
Glu Pro Leu Leu Val Ala Ala Gly Gly Gly Gly Arg Ala Tyr Leu
                230                 235                 240
Arg Pro Arg Asp Arg Gly Arg Thr Gln Ala Ser Pro Glu Lys Leu
                245                 250                 255
Glu Asn Arg Ser Glu Ala Pro Gly Ser Gly Gly Arg Gly Gly Ala
                260                 265                 270
Ala Gly Gly Gly Gly Gly Trp Thr Ser Arg Ala Pro Ser Pro Gln
                275                 280                 285
Ala Gly Arg Ser Leu Gln Glu Gly Ala Glu Gly Gly Gln Gly Cys
                290                 295                 300
Ser Glu Ala Trp Ala Thr Leu Gly Trp Ala Ala Ala Gly Gly Phe
                305                 310                 315
Gly Gly Gly Gly Gly Ala Cys Thr Ala Gly Gly Gly Gly Gly Gly
                320                 325                 330
Tyr Arg Gly Gly Asp Ala Ser Glu Thr Asp Asn Leu Trp Ala Asp
                335                 340                 345
Gly Glu Asp Gly Val Ser Phe Ile His Pro Ser Ser Glu Leu Phe
                350                 355                 360
Leu Gln Pro Leu Ala Val Thr Glu Asn His Gly Glu Val Glu Ile
                365                 370                 375
Arg Arg His Leu Asn Cys Ser His Cys Pro Leu Arg Asp Cys Gln
                380                 385                 390
Trp Gln Ala Glu Leu Gln Leu Ala Glu Cys Leu Cys Pro Glu Gly
                395                 400                 405
Met Glu Leu Ala Val Asp Asn Val Thr Cys Met Asp Leu His Lys
```

```
                410                 415                 420
Pro Pro Gly Pro Leu Val Leu Met Val Ala Val Val Ala Thr Ser
                425                 430                 435
Thr Leu Ser Leu Leu Met Val Cys Gly Val Leu Ile Leu Gly Gly
                440                 445                 450
Ala Trp Pro Gly Pro Val Leu Ala Ser Ala Thr Arg Cys His Arg
                455                 460                 465
Gly Phe Pro Ser Gln Cys Tyr Ser Ala Gln Ser Pro Gly Pro Trp
                470                 475                 480
Cys Leu Trp Gly Gly Val
                485

<210> 34
<211> 256
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509265CD1

<400> 34
Met Asp Gln Tyr Cys Ile Leu Gly Arg Ile Gly Glu Gly Ala His
  1               5                  10                  15
Gly Ile Val Phe Lys Ala Lys His Val Glu Pro Arg Val Gly Trp
                 20                  25                  30
Gln Cys Leu Pro Ser Ile Leu Gln Thr Gly Glu Ile Val Ala Leu
                 35                  40                  45
Lys Lys Val Ala Leu Arg Arg Leu Glu Asp Gly Phe Pro Asn Gln
                 50                  55                  60
Ala Leu Arg Glu Ile Lys Ala Leu Gln Glu Met Glu Asp Asn Gln
                 65                  70                  75
Tyr Val Val Gln Leu Lys Ala Val Phe Pro His Gly Gly Gly Phe
                 80                  85                  90
Val Leu Ala Phe Glu Phe Met Leu Ser Asp Leu Ala Glu Val Val
                 95                 100                 105
Arg His Ala Gln Arg Pro Leu Ala Gln Ala Gln Val Lys Ser Tyr
                110                 115                 120
Leu Gln Met Leu Leu Lys Gly Val Ala Phe Cys His Ala Asn Asn
                125                 130                 135
Ile Val His Arg Asp Leu Lys Pro Ala Asn Leu Leu Ile Ser Ala
                140                 145                 150
Ser Gly Gln Leu Lys Ile Ala Asp Phe Gly Leu Ala Arg Val Phe
                155                 160                 165
Ser Pro Asp Gly Ser Arg Leu Tyr Thr His Gln Val Ala Thr Arg
                170                 175                 180
Ser Ser Leu Ser Cys Arg Thr Thr Thr Arg Ser Pro Leu Arg Ser
                185                 190                 195
Arg Cys Pro Cys Pro Trp Arg Arg Cys Cys Leu Thr Ser Leu Pro
                200                 205                 210
Arg His Trp Ile Cys Trp Val Asn Ser Phe Ser Thr Leu Leu Thr
                215                 220                 225
Ser Ala Ser Gln Leu Pro Arg Leu Ser Ser Ile Ser Thr Ser Ser
                230                 235                 240
Gln Leu Pro Cys Leu Pro Ile His Leu Ser Cys Arg Phe Leu Ser
                245                 250                 255
```

Val

<210> 35
<211> 458
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509371CD1

<400> 35
```
Met Ser Thr Glu Gly Arg Leu Pro Ser Cys Ser Ala Cys Val Lys
  1               5                  10                  15
Gly Glu Leu Arg Val Leu Thr Ser Ala Ala Leu Thr Ser Arg Asp
                 20                  25                  30
Gly Pro Arg Pro Cys His Val Leu Phe Arg Ile Val His Leu Cys
                 35                  40                  45
Leu Arg Lys Ala Asp Gln Lys Leu Val Ile Ile Lys Gln Ile Pro
                 50                  55                  60
Val Glu Gln Met Thr Lys Glu Glu Arg Gln Ala Ala Gln Asn Glu
                 65                  70                  75
Cys Gln Val Leu Lys Leu Leu Asn His Pro Asn Val Ile Glu Tyr
                 80                  85                  90
Tyr Glu Asn Phe Leu Glu Asp Lys Ala Leu Met Ile Ala Met Glu
                 95                 100                 105
Tyr Ala Pro Gly Gly Thr Leu Ala Glu Phe Ile Gln Lys Arg Cys
                110                 115                 120
Asn Ser Leu Leu Glu Glu Glu Thr Ile Leu His Phe Phe Val Gln
                125                 130                 135
Ile Leu Leu Ala Leu His His Val His Thr His Leu Ile Leu His
                140                 145                 150
Arg Asp Leu Lys Thr Gln Asn Ile Leu Leu Asp Lys His Arg Met
                155                 160                 165
Val Val Lys Ile Gly Asp Phe Gly Ile Ser Lys Ile Leu Ser Ser
                170                 175                 180
Lys Ser Lys Ala Tyr Thr Val Val Gly Thr Pro Cys Tyr Ile Ser
                185                 190                 195
Pro Glu Leu Cys Glu Gly Lys Pro Tyr Asn Gln Lys Ser Asp Ile
                200                 205                 210
Trp Ala Leu Gly Cys Val Leu Tyr Glu Leu Ala Ser Leu Lys Arg
                215                 220                 225
Ala Phe Glu Ala Ala Asn Leu Pro Ala Leu Val Leu Lys Ile Met
                230                 235                 240
Ser Gly Thr Phe Ala Pro Ile Ser Asp Arg Tyr Ser Pro Glu Leu
                245                 250                 255
Arg Gln Leu Val Leu Ser Leu Leu Ser Leu Glu Pro Ala Gln Arg
                260                 265                 270
Pro Pro Leu Ser His Ile Met Ala Gln Pro Leu Cys Ile Arg Ala
                275                 280                 285
Leu Leu Asn Leu His Thr Asp Val Gly Ser Val Arg Met Arg Arg
                290                 295                 300
Ala Glu Lys Ser Val Ala Pro Ser Asn Thr Gly Ser Arg Thr Thr
                305                 310                 315
Ser Val Arg Cys Arg Gly Lys Trp Glu Glu Ala Ala Ser Pro His
```

```
                320                 325                 330
Gly Cys His Thr Ile Pro Ile Ser Leu Ile Val Pro Met His Cys
                335                 340                 345
Thr Cys Arg Tyr Pro Pro Gly Thr Cys Glu Ala Ser His Pro Thr
                350                 355                 360
Thr Thr Val Val Ser Val Cys Leu Gly Trp Trp Ala Gly His Pro
                365                 370                 375
Pro Ala Ala Ala Asn Ala Gln His Arg Gly Gly Pro Gly Gly Ser
                380                 385                 390
Trp Ala His Ala Glu Ser Arg Arg His Ala Leu Trp Ala Ser His
                395                 400                 405
Pro Val Gly Gly Pro Thr Pro Arg Cys Arg Arg Arg Gln Ser Pro
                410                 415                 420
Ser Trp Gly Ser Gly Ala Ala Thr Ala Pro Val His Leu Ala Phe
                425                 430                 435
Pro Gly Gly Pro Val Gly Cys Asp His Gln Ala Arg Gly Leu Trp
                440                 445                 450
Gly Leu Leu His Cys Leu Pro Asp
                455

<210> 36
<211> 362
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509389CD1

<400> 36
Met His Gln Ala Ala Leu Trp Leu Ser Ser Ser Lys Ser Ala Val
  1               5                  10                  15
Ile Pro Cys Trp Arg Arg Arg Pro Ser Cys Thr Ser Ser Cys Arg
                 20                  25                  30
Ser Cys Leu His Cys Ile Met Cys Thr Pro Thr Ser Ser Cys Thr
                 35                  40                  45
Glu Thr Ser Arg Pro Arg Thr Ser Cys Leu Thr Asn Thr Ala Trp
                 50                  55                  60
Ser Ser Arg Ser Val Val Gly Thr Pro Cys Tyr Ile Ser Pro Glu
                 65                  70                  75
Leu Cys Glu Gly Lys Pro Tyr Asn Gln Lys Ser Asp Ile Trp Ala
                 80                  85                  90
Leu Gly Cys Val Leu Tyr Glu Leu Ala Ser Leu Lys Arg Ala Phe
                 95                 100                 105
Glu Ala Ala Asn Leu Pro Ala Leu Val Leu Lys Ile Met Ser Gly
                110                 115                 120
Thr Phe Ala Pro Ile Ser Asp Arg Tyr Ser Pro Glu Leu Arg Gln
                125                 130                 135
Leu Val Leu Ser Leu Leu Ser Leu Glu Pro Ala Gln Arg Pro Pro
                140                 145                 150
Leu Ser His Ile Met Ala Gln Pro Leu Cys Ile Arg Ala Leu Leu
                155                 160                 165
Asn Leu His Thr Asp Val Gly Ser Val Arg Met Arg Arg Pro Val
                170                 175                 180
Gln Gly Gln Arg Ala Val Leu Gly Gly Arg Val Trp Ala Pro Ser
                185                 190                 195
```

```
Gly Ser Thr Gly Gly Leu Arg Gln Arg Glu Thr Trp Gly Lys Ser
                200                 205                 210
Ser Leu Pro Ala Cys Arg Asn Val Arg Arg Val Phe Val Leu Arg
                215                 220                 225
Pro Pro Ser Val Leu Gln Gly Arg Glu Val Arg Gly Pro Gln Gln
                230                 235                 240
His Arg Glu Gln Asp His Gln Cys Pro Leu Gln Arg Tyr Pro Pro
                245                 250                 255
Gly Thr Cys Glu Ala Ser His Pro Thr Thr Thr Val Val Ser Val
                260                 265                 270
Cys Leu Gly Trp Trp Ala Gly His Pro Pro Ala Ala Ala Asn Ala
                275                 280                 285
Gln His Arg Gly Gly Pro Gly Gly Ser Trp Ala His Ala Glu Ser
                290                 295                 300
Arg Arg His Ala Leu Trp Ala Ser His Pro Val Gly Gly Pro Thr
                305                 310                 315
Pro Arg Cys Arg Arg Arg Gln Ser Pro Ser Trp Gly Ser Gly Ala
                320                 325                 330
Ala Thr Ala Pro Val His Leu Ala Phe Pro Gly Gly Pro Val Gly
                335                 340                 345
Cys Asp His Gln Ala Arg Gly Leu Trp Gly Leu Leu His Cys Leu
                350                 355                 360
Pro Asp


<210> 37
<211> 292
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507005CD1

<400> 37
Met Cys Gln Arg Leu Trp Pro Trp Pro Ala Asn Gln Pro Leu Pro
  1               5                  10                  15
Gly Gly Leu Leu Pro Arg Pro Leu Ser Leu Ala Pro Ser Ser Ser
                 20                  25                  30
Ser Ser Cys Cys Ser Pro Pro Cys Ser Gln Asp Gly Arg Met Ala
                 35                  40                  45
Ala Gln Gly Ala Pro Arg Phe Leu Leu Thr Phe Asp Phe Asp Glu
                 50                  55                  60
Thr Ile Val Asp Glu Asn Ser Asp Asp Ser Ile Val Arg Ala Ala
                 65                  70                  75
Pro Gly Gln Arg Leu Pro Glu Ser Leu Arg Ala Thr Tyr Arg Glu
                 80                  85                  90
Gly Phe Tyr Asn Glu Tyr Met Gln Arg Val Phe Lys Tyr Leu Gly
                 95                 100                 105
Glu Gln Gly Val Arg Pro Arg Asp Leu Ser Ala Ile Tyr Glu Ala
                110                 115                 120
Ile Pro Leu Ser Pro Gly Met Ser Asp Leu Leu Gln Phe Val Ala
                125                 130                 135
Lys Gln Gly Ala Cys Phe Glu Val Ile Leu Ile Ser Asp Ala Asn
                140                 145                 150
Thr Phe Gly Val Glu Ser Ser Leu Arg Ala Ala Gly His His Ser
```

```
                155                 160                 165
Leu Phe Arg Arg Ile Leu Ser Asn Pro Ser Gly Pro Asp Ala Arg
                170                 175                 180
Gly Leu Leu Ala Leu Arg Pro Phe His Thr His Ser Cys Ala Arg
                185                 190                 195
Cys Pro Ala Asn Met Cys Lys His Lys Val Leu Ser Asp Tyr Leu
                200                 205                 210
Arg Glu Arg Ala His Asp Gly Val His Phe Glu Arg Leu Phe Tyr
                215                 220                 225
Val Gly Asp Gly Ala Asn Asp Phe Cys Pro Met Gly Leu Leu Ala
                230                 235                 240
Gly Gly Asp Val Ala Phe Pro Arg Arg Gly Tyr Pro Met His Arg
                245                 250                 255
Leu Ile Gln Glu Ala Gln Lys Ala Glu Pro Ser Ser Phe Arg Ala
                260                 265                 270
Ser Val Val Pro Trp Glu Thr Ala Ala Asp Val Arg Leu His Leu
                275                 280                 285
Gln Gln Val Leu Lys Ser Cys
                290

<210> 38
<211> 226
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509142CD1

<400> 38
Met Ser Ser Pro Arg Asp Phe Arg Ala Glu Pro Val Asn Asp Tyr
 1               5                  10                  15
Glu Gly Asn Asp Ser Glu Ala Glu Asp Leu Asn Phe Arg Glu Thr
                20                  25                  30
Leu Pro Ser Ser Ser Gln Glu Asn Thr Pro Arg Ser Lys Val Phe
                35                  40                  45
Glu Asn Lys Val Asn Ser Glu Lys Val Lys Leu Ser Leu Arg Asn
                50                  55                  60
Phe Pro His Asn Asp Tyr Glu Asp Val Phe Glu Glu Pro Ser Glu
                65                  70                  75
Ser Gly Ser Asp Pro Ser Met Trp Thr Ala Arg Gly Pro Phe Arg
                80                  85                  90
Arg Asp Arg Trp Ser Ser Glu Asp Glu Glu Ala Ala Gly Pro Ser
                95                 100                 105
Gln Ala Leu Ser Pro Leu Leu Ser Asp Thr Arg Lys Ile Val Ser
               110                 115                 120
Glu Gly Glu Leu Asp Gln Leu Ala Gln Ile Arg Pro Leu Ile Phe
               125                 130                 135
Asn Phe His Glu Gln Thr Ala Ile Lys Asp Cys Leu Lys Ile Leu
               140                 145                 150
Glu Glu Lys Thr Ala Ala Tyr Asp Ile Met Gln Glu Phe Met Ala
               155                 160                 165
Leu Glu Leu Lys Asn Leu Pro Gly Glu Phe Asn Ser Gly Asn Gln
               170                 175                 180
Pro Ser Asn Arg Glu Lys Asn Arg Tyr Arg Asp Ile Leu Pro Phe
               185                 190                 195
```

```
Gln His His Gly Tyr Ser Gly Pro Asn Glu Arg Thr Thr Phe Trp
                200                 205                 210
His Gly Ser Asn Glu Gly Ala Val Ser Leu Leu Leu Arg Tyr Cys
                215                 220                 225
Ala


<210> 39
<211> 261
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509157CD1

<400> 39
Met Gln Glu Phe Met Ala Leu Glu Leu Lys Asn Leu Pro Gly Glu
 1                5                  10                  15
Phe Asn Ser Gly Asn Gln Pro Ser Asn Arg Glu Lys Asn Arg Tyr
                 20                  25                  30
Arg Asp Ile Leu Pro Tyr Asp Ser Thr Arg Val Pro Leu Gly Lys
                 35                  40                  45
Ser Lys Asp Tyr Ile Asn Ala Ser Tyr Ile Arg Ile Val Asn Cys
                 50                  55                  60
Gly Glu Glu Tyr Phe Tyr Ile Ala Thr Gln Gly Pro Leu Leu Ser
                 65                  70                  75
Thr Ile Asp Asp Phe Trp Gln Met Val Leu Glu Asn Asn Ser Asn
                 80                  85                  90
Val Ile Ala Met Ile Thr Arg Glu Ile Glu Gly Gly Ile Ile Lys
                 95                 100                 105
Cys Tyr His Tyr Trp Pro Ile Ser Leu Lys Lys Pro Leu Glu Leu
                110                 115                 120
Lys His Phe Arg Val Phe Leu Glu Asn Tyr Gln Ile Leu Gln Tyr
                125                 130                 135
Phe Ile Ile Arg Met Phe Gln Val Val Glu Lys Ser Thr Gly Thr
                140                 145                 150
Ser His Ser Val Lys Gln Leu Gln Phe Thr Lys Trp Pro Asp His
                155                 160                 165
Gly Thr Pro Ala Ser Ala Asp Ser Phe Ile Lys Tyr Ile Arg Tyr
                170                 175                 180
Ala Arg Lys Ser His Leu Thr Gly Pro Met Val Val His Cys Ser
                185                 190                 195
Ala Gly Ile Gly Arg Thr Gly Val Phe Leu Cys Val Asp Val Val
                200                 205                 210
Phe Cys Ala Ile Val Lys Asn Cys Ser Phe Asn Ile Met Asp Ile
                215                 220                 225
Val Ala Gln Met Arg Glu Gln Arg Ser Gly Met Val Gln Thr Lys
                230                 235                 240
Glu Gln Tyr His Phe Cys Tyr Asp Ile Val Leu Glu Val Leu Arg
                245                 250                 255
Lys Leu Leu Thr Leu Asp
                260

<210> 40
<211> 173
```

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509246CD1

<400> 40
Met Pro Cys Lys Ser Ala Glu Trp Leu Gln Glu Glu Leu Glu Ala
1 5 10 15
Arg Gly Gly Ala Ser Leu Leu Leu Leu Asp Cys Arg Pro His Glu
20 25 30
Leu Phe Glu Ser Ser His Ile Glu Thr Ala Ile Asn Leu Ala Ile
35 40 45
Pro Gly Leu Met Leu Arg Arg Leu Arg Lys Gly Asn Leu Pro Ile
50 55 60
Arg Ser Ile Ile Pro Asn His Ala Asp Lys Glu Arg Phe Ala Thr
65 70 75
Arg Cys Lys Ala Ala Thr Val Leu Leu Tyr Asp Glu Ala Thr Ala
80 85 90
Glu Trp Gln Pro Glu Pro Gly Ala Pro Ala Ser Val Leu Gly Leu
95 100 105
Leu Leu Gln Lys Leu Arg Asp Asp Gly Cys Gln Ala Tyr Tyr Leu
110 115 120
Gln Asp Pro Ala Leu Pro Leu Pro Arg Leu Arg Gln Gly Leu His
125 130 135
Gln Pro Gly Arg Ala Arg Gln Val Trp His Gln Val Tyr Pro Gln
140 145 150
Cys His Thr Gln Pro Thr Gln Arg Leu Arg Ala Arg Arg Arg Val
155 160 165
His Leu Gln Ala Asp Pro His Leu
170

<210> 41
<211> 412
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509380CD1

<400> 41
Met Ser Ser Pro Arg Asp Phe Arg Ala Glu Pro Val Asn Asp Tyr
1 5 10 15
Glu Gly Asn Asp Ser Glu Ala Glu Asp Leu Asn Phe Arg Glu Thr
20 25 30
Leu Pro Ser Ser Ser Gln Glu Asn Thr Pro Arg Ser Lys Val Phe
35 40 45
Glu Asn Lys Val Asn Ser Glu Lys Val Lys Leu Ser Leu Arg Asn
50 55 60
Phe Pro His Asn Asp Tyr Glu Asp Val Phe Glu Glu Pro Ser Glu
65 70 75
Ser Gly Ser Asp Pro Ser Met Trp Thr Ala Arg Gly Pro Phe Arg
80 85 90
Arg Asp Arg Trp Ser Ser Glu Asp Glu Glu Ala Ala Gly Pro Ser

```
                    95                  100                 105
Gln Ala Leu Ser Pro Leu Leu Ser Asp Thr Arg Lys Ile Val Ser
                110                 115                 120
Glu Gly Glu Leu Asp Gln Leu Ala Gln Ile Arg Pro Leu Ile Phe
                125                 130                 135
Asn Phe His Glu Gln Thr Ala Ile Lys Asp Cys Leu Lys Ile Leu
                140                 145                 150
Glu Glu Lys Thr Ala Ala Tyr Asp Ile Met Gln Glu Phe Met Ala
                155                 160                 165
Leu Glu Leu Lys Asn Leu Pro Gly Glu Phe Asn Ser Gly Asn Gln
                170                 175                 180
Pro Ser Asn Arg Glu Lys Asn Arg Tyr Arg Asp Ile Leu Pro Tyr
                185                 190                 195
Asp Ser Thr Arg Val Pro Leu Gly Lys Ser Lys Asp Tyr Ile Asn
                200                 205                 210
Ala Ser Tyr Ile Arg Ile Val Asn Cys Gly Glu Glu Tyr Phe Tyr
                215                 220                 225
Ile Ala Thr Gln Gly Pro Leu Leu Ser Thr Ile Asp Asp Phe Trp
                230                 235                 240
Gln Met Val Leu Glu Asn Asn Ser Asn Val Ile Ala Met Ile Thr
                245                 250                 255
Arg Glu Ile Glu Gly Gly Ile Ile Lys Cys Tyr His Tyr Trp Pro
                260                 265                 270
Ile Ser Leu Lys Lys Pro Leu Glu Leu Lys His Phe Arg Val Phe
                275                 280                 285
Leu Glu Asn Tyr Gln Ile Leu Gln Tyr Phe Ile Ile Arg Met Phe
                290                 295                 300
Gln Val Val Glu Lys Ser Thr Gly Thr Ser His Ser Val Lys Gln
                305                 310                 315
Leu Gln Phe Thr Lys Trp Pro Asp His Gly Thr Pro Ala Ser Ala
                320                 325                 330
Asp Ser Phe Ile Lys Tyr Ile Arg Tyr Ala Arg Lys Ser His Leu
                335                 340                 345
Thr Gly Pro Met Val Val His Cys Ser Ala Gly Ile Gly Arg Thr
                350                 355                 360
Gly Val Phe Leu Cys Val Asp Val Val Phe Cys Ala Ile Val Lys
                365                 370                 375
Asn Cys Ser Phe Asn Ile Met Asp Ile Val Ala Gln Met Arg Glu
                380                 385                 390
Gln Arg Ser Gly Met Val Gln Thr Lys Glu Leu Ser Ser Val Ala
                395                 400                 405
Glu Asn Tyr Lys Cys Val Phe
                410
```

```
<210> 42
<211> 197
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509382CD1

<400> 42
Met Ser Ser Pro Arg Asp Phe Arg Ala Glu Pro Val Asn Asp Tyr
  1               5                  10                  15
```

```
Glu Gly Asn Asp Ser Glu Ala Glu Asp Leu Asn Phe Arg Glu Thr
                20                  25                  30
Leu Pro Ser Ser Ser Gln Glu Asn Thr Pro Arg Ser Lys Val Phe
                35                  40                  45
Glu Asn Lys Val Asn Ser Glu Lys Val Lys Leu Ser Leu Arg Asn
                50                  55                  60
Phe Pro His Asn Asp Tyr Glu Asp Val Phe Glu Glu Pro Ser Glu
                65                  70                  75
Ser Gly Ser Asp Pro Ser Met Trp Thr Ala Arg Gly Pro Phe Arg
                80                  85                  90
Arg Asp Arg Trp Ser Ser Glu Asp Glu Glu Ala Ala Gly Pro Ser
                95                  100                 105
Gln Ala Leu Ser Pro Leu Leu Ser Asp Thr Arg Lys Ile Val Ser
                110                 115                 120
Glu Gly Glu Leu Asp Gln Leu Ala Gln Ile Arg Pro Leu Ile Phe
                125                 130                 135
Asn Phe His Glu Gln Thr Ala Ile Lys Asp Cys Leu Lys Ile Leu
                140                 145                 150
Glu Glu Lys Thr Ala Ala Tyr Asp Ile Met Gln Glu Phe Met Ala
                155                 160                 165
Leu Glu Leu Lys Asn Leu Pro Gly Glu Phe Asn Ser Gly Asn Gln
                170                 175                 180
Pro Ser Asn Arg Glu Lys Asn Arg Tyr Arg Asp Ile Leu Pro Tyr
                185                 190                 195
Gly Asn


<210> 43
<211> 3853
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503679CB1

<400> 43
cggctcgagc acatgcatgt ctttattgta ggcatgagcc tgtcctctgt gacgctggcc 60
agcgccctac aggtcagggg tgaagctctg tctgaggagg aaatctggtc cctcctgttc 120
ctggccgctg agcagctcct ggaagacctc cgcaacgatt cctcggacta tgtggtttgc 180
ccctggtcag ccctgctttc tgcagctgga agcctttctt tccaaggccg tgtttctcat 240
atagaggctg ctcctttcaa ggcccctgaa ctgctacagg gacagagtga ggatgagcag 300
cctgatgcat ctcagcccct gcagctctgc gagcccctgc actccatcct gctgaccatg 360
tgtgaagacc agcctcacag gcggtgcacg ttgcagtcgg ttctggaagc ttgtcgggtt 420
catgagaaag aagtgtctgt ctacccagcc cctgctggtc tccacatcag aaggctggtt 480
ggcttggttc tgggtaccat ttctgaggtg gagaaaagag ttgtggagga aagctcctct 540
gtgcagcaga acagaagcta cctgctcagg aagaggctgc gtgggacaag cagcgagagc 600
ccagcggcac aggccccgga gtgtctgcat ccttgcagag tttcagaaag aagcacggag 660
acccagagct caccagagcc ccattggagc accttgacac acagtcactg cagcctcctt 720
gttaaccgcg ctcttccagg agcagatccc caggaccagc aggcgggccg gaggctcagc 780
tctggatctg tgcactcggc agcagacagc tcatggccaa caactccttc tcagaggggt 840
tttctgcaaa gaaggagcaa gttttccagg ccagagttca tcctgttggc tggagaggcc 900
ccgatgacac tacatctgcc gggatcggtt gtgaccaaaa aagggaaatc ctatttggct 960
ctcagggacc tctgtgtggt cctgctgaac gggcagcacc tggaggtaaa atgtgatgtt 1020
gaatcaacag tgggagctgt cttcaatgcc gtgacatcct ttgccaacct cgaggaactc 1080
acctactttg gcttggcgta tatgaaaagc aaagagttct ttttcctgga cagtgaaacc 1140
```

```
agattgtgca aaatagctcc tgaaggctgg agagagcagc ctcagaagac ctccatgaat 1200
accttcacac tcttcctgag gataaagttc tttgtcagcc actatgggct gctccagcac 1260
agcctgacaa ggcaccagtt ttacctgcag cttcggaaag atatcctgga ggagaggctg 1320
tactgcaatg aagagatact gctgcagctg gggtccttg ccttgcaggc tgagtttggc 1380
aattacccta aggagcaggt ggagagtaag ccatactttc acgttgaaga ttacatccca 1440
gcgagtctga tcgagaggat gaccgctcta cgggtccagg ttgaagtctc agagatgcac 1500
cggctcagct ctgcactgtg gggagaggat gctgagctga agttcttgag ggtcactcag 1560
cagctcccag aatacggtgt gctggttcac caagtattct cagagaagag gaggccagaa 1620
gaggagatgg ccctggggat ctgtgccaag ggtgtcatag tctatgaagt gaaaaacaac 1680
agcagaattg caatgttacg gtttcagtgg agagaaaccg ggaagatttc tacttatcaa 1740
aaaaagttca ccatcacaag cagtgtcact gggaagaagc acacatttgt cacagattca 1800
gccaagacca gtaaatactt actggacctc tgctcagccc agcatgggtt taatgcacag 1860
atgggctctg ggcagccttc ccatgtttta tttgaccatg ataagtttgt gcaaatggcc 1920
aatttgagtc ctgcacacca ggcccggtct aagcctctca tttggattca gagattgtca 1980
tgctcagaaa acgagttgtt tgtatccagg cttcagggtg ctgcaggagg cctgctgagt 2040
acatcaatgg ataacttcaa cgtggacggc agcaaggagg ctggagcaga aggcatcggg 2100
cgcagcccct gcactggccg ggagcagctg aagagtgcct gtgtgatcca gaagccaatg 2160
acctgggact ctctctctgg accacctgtt cagagcatgc atgcaggtgt tggtggaaat 2220
aacccagatg aagaaaagaa tggcacagcc aattctgggg tctcctctac agacatcctg 2280
agcttcgggt accagggaag tttgttgtca cacacacaag accaggacag aaatactgaa 2340
gaactagaca tggctggggt gcagagctta gtgcccaggc tgagacatca gctttccttt 2400
ctgccgttaa agggtgctgg ttcttcttgt cctccatcac ctccagaaat cagtgctggt 2460
gaaatctact ttgtggaact ggttaaagaa gatgggacac ttggattcag tgtaactggt 2520
ggcattaaca ccagtgtgcc atatggtggt atctatgtga aatccattgt tcctggagga 2580
ccagctgcca aggaagggca gatcctacag ggtgaccgac tcctgcaggt ggatggagtg 2640
attctgtgcg gcctcaccca caagcaggct gtgcagtgcc tgaagggtcc tgggcaggtt 2700
gcaagactgg tcttagagag aagagtcccc aggagtacac agcagtgtcc ttctgctaat 2760
gacagcatgg gagatgaacg cacggctgtt tccttggtaa cagccttgcc tggcaggcct 2820
tcgagctgtg tctcggtgac agatggtcct aagtttgaag tcaaactaaa aaagaatgcc 2880
aatggtttgg gattcagttt cgtgcagatg gagaaagaga gctgcagcca tctcaaaagt 2940
gatcttgtga ggattaagag gctctttccg gggcagccag ctgaggagaa tggggccatt 3000
gcagctggtg acattatcct ggccgtgaat ggaaggtcca cggaaggcct catcttccag 3060
gaggtgctgc atttactgag aggggcccca caggaagtca cgctcctcct ttgccgaccc 3120
cctccaggtg cgctgcctga gctggagcag gaatggcaga cacctgaact ctcagctgac 3180
aaagaattca ccagggcaac atgtactgac tcatgtacca gccccatcct ggatcaagag 3240
gacagctgga gggacagtgc ctccccagat gcaggggaag gcctgggtct caggccagag 3300
tcttcccaaa aggccatcag agaggcacaa tggggccaaa acagagagag accttgggcc 3360
agttccttga cacattctcc tgagtcccac cctcatttat gcaaacttca ccaagaaagg 3420
gatgaatcaa cattggcgac ctctttggaa aaggatgtga ggcaaaactg ctattcagtt 3480
tgtgatatca tgagacttgg aagatattcc ttctcatctc ctctaaccag actttcgaca 3540
gatattttct gagcaccttc tctgcatgtc tgcagtgctg tgtaaaatgc cctacctttg 3600
catggactat tctttctaat caagaggcgt gtgtggcgaa cttggggcag cccctggaag 3660
tcttgttctt tgaccattac gtctgcggct gcatcaccag ataatgagct tcaccactcg 3720
tctgcctcct gtgtccttcc gcggggagta aatgtcactt cagcttgccg catctctaaa 3780
taggcaaatt ttcagtgctc agaaaaggac ctgatctttg cacaaagtgc tttgatggtt 3840
gcctgcttgg ggc                                                    3853
```

<210> 44
<211> 4000
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503681CB1

```
<400> 44
tgcatgtctt tattgtaggc atgagcctgt cctctgtgac gctggccagc gccctacagg 60
tcaggggtga agctctgtct gaggaggaaa tctggtccct cctgttcctg gccgctgagc 120
agctcctgga agacctccgc aacgcccctg cagctctgcg agcccctgca ctccatcctg 180
ctgaccatgt gtgaagacca gcctcacagg cggtgcacgt tgcagtcggt tctggaagct 240
tgtcgggttc atgagaaaga agtgtctgtc tacccagccc ctgctggtct ccacatcaga 300
aggctggttg gcttggttct gggtaccatt tctgaggtgg agaaaagagt tgtggaggaa 360
agctcctctg tgcagcagaa cagaagctac ctgctcagga agaggctgcg tgggacaagc 420
agcgagagcc cagcggcaca ggcccccggag tgtctgcatc cttgcagagt ttcagaaaga 480
agcacggaga cccagagctc accagagccc cattggagca ccttgacaca cagtcactgc 540
agcctccttg ttaaccgcgc tcttccagga gcagatcccc aggaccagca ggcgggccgg 600
aggctcagct ctggatctgt gcactcggca gcagacagct catggccaac aactccttct 660
cagaggggtt ttctgcaaag aaggagcaag ttttccaggc cagagttcat cctgttggct 720
ggagaggccc cgatgacact acatctgccg ggatcggttg tgaccaaaaa agggaaatcc 780
tatttggctc tcagggacct ctgtgtggtc ctgctgaacg ggcagcacct ggaggtaaaa 840
tgtgatgttg aatcaacagt gggagctgtc ttcaatgccg tgacatcctt tgccaacctc 900
gaggaactca cctactttgg cttggcgtat atgaaaagca aagagttctt tttcctggac 960
agtgaaacca gattgtgcaa aatagctcct gaaggctgga gagagcagcc tcagaagacc 1020
tccatgaata ccttcacact cttcctgagg ataaagttct ttgtcagcca ctatgggctg 1080
ctccagcaca gcctgacaag gcaccagttt tacctgcagc ttcggaaaga tatcctggag 1140
gagaggctgt actgcaatga agagatactg ctgcagctgg gggtccttgc cttgcaggct 1200
gagtttggca attaccctaa ggagcaggtg gagagtaagc catactttca cgttgaagat 1260
tacatcccag cgagtctgat cgagaggatg accgctctac gggtccaggt tgaagtctca 1320
gagatgcacc ggctcagctc tgcactgtgg ggagaggatg ctgagctgaa gttcttgagg 1380
gtcactcagc agctcccaga atacggtgtg ctggttcacc aagtattctc agagaagagg 1440
aggccagaag aggagatggc cctggggatc tgtgccaagg gtgtcatagt ctatgaagtg 1500
aaaaacaaca gcagaattgc aatgttacgg tttcagtgga gagaaaccgg gaagatttct 1560
acttatcaaa aaaagttcac catcacaagc agtgtcactg ggaagaagca cacatttgtc 1620
acagattcag ccaagaccag taaatactta ctggacctct gctcagccca gcatgggttt 1680
aatgcacaga tgggctctgg gcagccttcc catgttttat ttgaccatga taagtttgtg 1740
caaatggcca atttgagtcc tgcacaccag gcccggtcta agcctctcat ttggattcag 1800
agattgtcat gctcagaaaa cgagttgttt gtatccaggc ttcagggtgc tgcaggaggc 1860
ctgctgagta catcaatgga taacttcaac gtggacggca gcaaggaggc tggagcagaa 1920
ggcatcgggc gcagcccctg cactggccgg gagcagctga agagtgcctg tgtgatccag 1980
aagccaatga cctgggactc tctctctgga ccacctgttc agagcatgca tgcaggctca 2040
aagaataata ggaggaagag ctttatagct gaaccgggcc gagaaattgt acgtgtgaca 2100
ctgaaacgtg acccacatcg tggttttggg tttgtcatta atgagggaga gtattcaggc 2160
caagctgacc ctggcatttt tatatcttct attatacctg gaggaccagc agaaaaagca 2220
aaaacgatca aaccaggagg gcagatacta gccctgaatc acatcagtct ggagggcttc 2280
acattcaaca tggctgttag gatgatccag aattcccctg acaacataga attaattatt 2340
tctcagtcaa aaggtgttgg tggaaataac ccagatgaag aaaagaatag cacagccaat 2400
tctggggtct cctctacaga cgtcctgagc ttcgggtacc agggaagttt gttgtcacac 2460
acacaagacc aggacagaaa tactgaagaa ctagacatgg ctggggtgca gagcttagtg 2520
cccaggctga gacatcagct ttcctttctg ccgttaaagg gtgctggttc ttcttgtcct 2580
ccatcacctc cagaaatcag tgctggtgaa atctactttg tggaactggt taaagaagat 2640
gggacacttg gattcagtgt aactggtggc attaacacca gtgtgccata tggtggtatc 2700
tatgtgaaat ccattgttcc tggaggacca gctgccaagg aagggcagat cctacagggt 2760
gaccgactcc tgcaggtgga tggagtgatt ctgtgcggcc tcacccacaa gcaggctgtg 2820
cagtgcctga agggtcctgg gcaggttgca agactggtct tagagagaag agtccccagg 2880
agtacacagc agtgtccttc tgctaatgac agcatgggag atgaacgcac ggctgtttcc 2940
ttggtaacag ccttgcctgg caggccttcg agctgtgtct cggtgacaga tggtcctaag 3000
tttgaagtca aactaaaaaa gaatgccaat ggtttgggat tcagtttcgt gcagatggag 3060
aaagagagct gcagccatct caaaagtgat cttgtgagga ttaagaggct ctttccgggg 3120
cagccagctg aggagaatgg ggccattgca gctggtgaca ttatcctggc cgtgaatgga 3180
aggtccacgg aaggcctcat cttccaggag gtgctgcatt tactgagagg ggccccacag 3240
gaagtcacgc tcctcctttg ccgacccccct ccaggtgcgc tgcctgagct ggagcaggaa 3300
```

tggcagacac ctgaactctc agctgacaaa gaattcacca gggcaacatg tactgactca 3360
tgtaccagcc ccatcctgga tcaagaggac agctggaggg acagtgcctc cccagatgca 3420
ggggaaggcc tgggtctcag gccagagtct tcccaaaagg ccatcagaga ggcacaatgg 3480
ggccaaaaca gagagagacc ttgggccagt tccttgacac attctcctga gtcccaccct 3540
catttatgca aacttcacca agaaagggat gaatcaacat tggcgacctc tttggaaaag 3600
gatgtgaggc aaaactgcta ttcagtttgt gatatcatga gacttggaag atattccttc 3660
tcatctcctc taaccagact ttcgacagat attttctgag caccttctct gcatgtctgc 3720
agtgctgtgt aaaatgccct acctttgcat ggactattct ttctaatcaa gaggcgtgtg 3780
tggcgaactt ggggcagccc ctggaagtct tgttctttga ccattacgtc tgcggctgca 3840
tcaccagata atgagcttca ccactcgtct gcctcctgtg tccttccgcg gggagtaaat 3900
gtcacttcag cttgccgcat ctctaaatag gcaaattttc agtgctcaga aaaggacctg 3960
atctttgcac aaagtgcttt gatggttgcc tgcttggggc 4000

<210> 45
<211> 6408
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505819CB1

<400> 45
gccttgtctc cggaaggact ctctcagaat gttcctcaaa tgtgacatgt cgattcacga 60
tgtttcggtg agtgcagctg agacacaggc gattgtagat gaggccttgg ggctgcgaaa 120
aaagaggcaa gcgctgattg tgcgggagaa ggagccggac ctgaaacttg tgcagcccat 180
tcctttcttc acctggaagt gcctcggaga gagcttgctg gccatgtaca atcatctcac 240
cacctgtgag cccccacgtc ccagccttgg caaaaggatt gatttgtcgg actaccagga 300
ccccagccag cctcttgagt cctccatggt ggtgacgcca gttaacgtga tccagccaag 360
cactgtcagc accaacccag ctgtggctgt cgccgagcct gtggtctcct acacctctgt 420
ggctacaacc agcttcccac tgcacagtcc tggtctgttg gagacaggcg ctcctgtggg 480
tgatatttct gggggagata aatccaagaa aggggtaaaa cggaagaaga tttcagaaga 540
gagtggagaa acagcaaagc ggcggtctgc ccgtgtccga aacaccaagt gcaaaaaaga 600
agagaaagta gacttccagg agcttctgat gaagttcttg ccgtccaggt taagaaagct 660
ggaccctgag gaggaagatg attcctttaa taactatgaa gtccagtcag aagccaaact 720
ggaaagcttc ccaagcattg ggcctcaaag gctgtcattt gactcagcca cattcatgga 780
atctgaaaag caggacgtgc atgagttcct gctggagaac ctaaccaacg ggggcatcct 840
ggagctgatg atgcgctacc tgaaagccat gggccacaag ttcttggtaa ggtggcctcc 900
aggcttggcg gaggtcgtgc tcagcgtcta ccacagctgg aggaggcaca gcaccagcct 960
gcccaacccg ctgctgaggg actgcagcaa caagcacatc aaggacagga tgctgatgtc 1020
tctctcctgc atggaactcc agctggacca gtggctgctg accaaaggca gaagctctgc 1080
agtgtctcct cggaactgcc ctgctggtat ggtgaatggc agatttggac ctgacttccc 1140
agggacccac tgcctgggtg acctcctaca gctgtcattt gcctcgtccc agcgcgacct 1200
gttcgaggat ggttggctgg agtttgtggt ccgtgtttac tggctgaagg ctcgcttcct 1260
ggcgctgcag ggagacatgg agcaggccct ggagaactat gacatctgca cagaaatgct 1320
ccagagttcc accgccatcc aggtggaggc aggggctgaa cgaagagaca ttgtcatccg 1380
gctgcccaac ctccataatg actctgtggt ttccctggag gagattgata agaacctgaa 1440
gtcgctggag cggtgccagt ccctggagga gattcagcgg ctgtatgaag caggcgacta 1500
caaggctgtt gtgcatctgc tccgccccac tttgtgcacc agtgggtttg accgggccaa 1560
acacctggag tttatgactt ccattcctga gaggccagcc cagctgcttc ttctgcagga 1620
ctccttgctc cggctgaagg actatcggca gtgttttgag tgttccgatg tggctctgaa 1680
cgaggctgtc cagcagatgg tgaactcagg tgaggctgcc gccaaggagg agtgggtggc 1740
cacagtgacc caactgctga tgggcatcga gcaggccctc tctgcggaca gcagtggtag 1800
catcctgaag gtatcatcct ccaccactgg ccttgtgcgg ctcaccaaca acctcatcca 1860
ggtcattgac tgcagcatgg ctgtgcagga ggaggccaag gagccccacg tctcttcagt 1920
gctaccctgg atcattctac accggatcat ctggcaggag gaagacacct tccattctct 1980

```
gtgccaccag cagcagctcc aaaacccagc ggaggaaggg atgtcagaga cgcccatgct 2040
cccatcctcc ctcatgctgc tgaacacagc ccacgagtat ttgggcagaa ggtcctggtg 2100
ctgcaattca gatggggctc tgctgcgatt ctatgtgcga gtactccaga aggaactggc 2160
tgcatccacc tctgaagaca cgcaccctta caaggaggag ctggagacag ccttggagca 2220
gtgcttctac tgcctgtaca gcttccccag caagaagagt aaggccaggt acctggagga 2280
acactcggcc cagcaggtgg atcttatatg ggaggatgca ctgttcatgt ttgagtattt 2340
taagcccaag acccttcctg aatttgacag ctataagacc agcaccgtgt ctgctgactt 2400
ggccaaccta ctgaagagaa ttgccaccat tgtgcctcgc acagagaggc cagcccttag 2460
cctggacaaa gtctctgcct acattgaggg aacttcaact gaggtaccct gcctcccaga 2520
gggggctgac ccctcccctc cagtggtgaa cgagctttac tacctcctgg ctgattatca 2580
tttcaaaaac aaggagcagt ccaaggccat caagttctac atgcatgaca tctgcatctg 2640
ccccaatagg tttgattcct gggcaggcat ggctctggcc cgggccagcc gcattcagga 2700
caagctgaac tccaatgagc tgaagagtga tgggcccatt tggaagcatg ccacgcccgt 2760
cttgaactgc ttccgtcggg ccctggagat tgacagctcc aacttgtccc tatggattga 2820
gtatggcacc atgtcctatg ccttgcactc attcgcctca cgtcaattga agcagtggag 2880
aggcgagctg ccccctgagc tcgtgcagca gatggagggc cggcgcgaca gcatgctaga 2940
gacagccaag cactgtttca catcagcagc ccgctgcgag ggtgatggtg acgaggagga 3000
gtggctcatc cactacatgc tgggcaaggt ggctgagaag cagcagcagc cacccaccgt 3060
ttacttgctg cactacaggc aggctggcca ctacctgcac gaggaggctg cccgctaccc 3120
caagaagatc cactaccaca acccacctga gctggccatg gaggccctgg aggtgtactt 3180
tcggctccat gcttccatcc tgaagctcct ggggaagccc gattctgggg ttggtgcaga 3240
ggtcctggtc aactttatga aggaggctgc agaaggaccc tttgccaggg gcgaggagaa 3300
gaacacaccc aaagcttcag aaaaggagaa ggcctgcctg gtggacgagg actcccactc 3360
ttcagctggg acactgccgg gccccggagc ctccctcccc tcctcctctg gcccaggtct 3420
gacatcccca ccttacacag ccactccgat tgaccacgat tacgtcaaat gtaaaaaacc 3480
ccaccagcag gcaacgccgg acgaccgaag ccaggacagc acagccgtag cactctcaga 3540
ctctagctca acgcaggact tctttaatga gcccaccagc ttactggaag gctccaggaa 3600
atcctacaca gagaagaggc tgcccattct cagttcccaa gcaggagcga cgggtaaaga 3660
tcttcagggg gccacagaag aaagaggaaa aaacgaggag tcattggaga gtacagaagg 3720
cttccgggct gcagagcaag gtgtccagaa gcctgctgca gaaaccccag cctctgcttg 3780
catccctggc aagccctcag catccacacc caccctgtgg gatgggaaga agagagggga 3840
cctcccaggg gagccagtgg ccttccccca ggggctgccg gctggtgctg aggagcagcg 3900
gcagtttctc acagagcagt gcatcgcctc cttccgcctg tgcctgagcc gcttccccca 3960
gcactataag agtctctacc gtctggcctt cctctacacc tacagcaaga cccaccggaa 4020
cctccagtgg gcccgcgacg tgttgctagg cagcagtatc ccgtggcaac aactgcagca 4080
catgccggca caggggctct tctgcgagag gaacaagacc aatttcttca acggcatctg 4140
gcggatcccc gtggacgaga ttgaccggcc gggcagcttt gcctggcaca tgaaccgctc 4200
catcgtgctg ctgctcaagg tgctggccca gctgcgggac cacagcaccc tgctgaaggt 4260
gtcctccatg cttcagcgga ccccagacca gggcaagaag tatctgcgag atgctgaccg 4320
ccaggtcctg gcgcagcggg ccttcatcct cactgtgaag gtgctcgaag acacgctgag 4380
cgagctcgca gaggggtcag aacgcccagg gcccaaggtc tgtggcctcc ccggagccag 4440
gatgaccacc gatgtctcac acaaggccag tcctgaggat ggccaggagg gcctccccca 4500
gccgaagaag ccccctctgg ctgatggctc agggccaggg cccgagccag gaggcaaagt 4560
gggcctcctc aaccaccggc ctgtggccat ggatgcagga gacagtgcag accaaagcgg 4620
ggagcggaag gataaagaga gcccacgggc agggcccact gagcccatgg acacgagtga 4680
ggccactgtt tgccactcag acttggagcg gacaccaccc ctgctgccag gtcgccccgc 4740
aagggaccgg ggccccgaga gccggcccac tgagctgtcc ctggaggagc tgagcatcag 4800
tgcccggcag cagcccaccc cgctcacccc agcccagcca gcccccgccc ccgcccccgc 4860
caccaccaca gggaccaggg caggggccca cccggaggag ccgctctccc ggctcagccg 4920
caagaggaag ctcctggagg acacagagtc aggcaagaca cttctgttgg atgcctaccg 4980
tgtgtggcag cagggccaga agggtgtggc ctatgacctg ggccgtgtgg agaggatcat 5040
gtcggagacc tacatgctca tcaagcaggt ggatgaggag gctgcgctgg agcaggctgt 5100
gaagttctgc caggtccatc ttggggctgc cgcccagaga caggcctcgg gggacacccc 5160
caccactcca aagcacccca aagacagccg agagaacttc tttcctgtga cagtggtgcc 5220
cacagcccct gaccctgtgc cagctgactc tgtccagcgg cccagtgatg ctcacaccaa 5280
gcctcgccct gcactagctg ccgccacaac tattatcacc tgccctccgt cagcatcagc 5340
```

ttccaccctg gaccagtcca aggaccctgg gcctccccgg ccacacaggc ctgaagctac 5400
ccccagcatg gcctctctgg gcccagaggg agaagagctg gcgagagtgg cagagggcac 5460
cagcttcccg cctcaggagc cacggcacag tccgcaggtg aagatggccc ccacaagttc 5520
cccggcagag ccacactgct ggccggcaga ggctgccctg ggcacaggcg ctgagcccac 5580
ctgcagccag gctgccagct ccaaggcccc cagcagtggg agtgcccagc caccagaggg 5640
tcacccaggc aagcctgagc ccagccgggc taagtcccgc cccctgccca acatgccaaa 5700
gctggtcatc cctccgccg ccaccaagtt ccccccctgag atcaccgtca cgccacccac 5760
cccaaccctg ctctccccca aaggcagcat ctcggaggag accaagcaga agctgaagtc 5820
agccatcctt tctgcccagt ctgctgccaa cgtgaggaag gagagcctat gccagccagc 5880
cctggaggtc ctggagacat ccagccagga gtcctcgctg gagagcgaga cagacgagga 5940
cgacgactac atggacattt gaggggccac tgcagcccca ccgccacgcc ccaggggacc 6000
agccaggcct ggaatgcccc ctgggcagga ccctgggcag gaccagaggc ccacatggat 6060
gccactcccc acacagcccc caggcctgcc cagcccacct cctcatggca tcctccctgt 6120
acccaggtca ggctgtccac accacatggg agcccagagg aggaggggcc cgccttagcc 6180
atgtgaaggt ggattggtcg ccatctgcac gccaggcggc atccttttct atgaagtgtt 6240
gactttgtaa atctgcccac acccagctgg ccatatccac ccctcgacgc cgggatgagc 6300
cggctctgcc tgtgtcacag tggaggggtc ctttagggcc aggctcaccc ctcacccttt 6360
ttttggttgc ttttctaata aagatggaac agttaaaaaa aaaaaaaa 6408

<210> 46
<211> 1072
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505083CB1

<400> 46
gcgatcttcc cagcacagac gtttggacag agcaggctcc taaggtctcc agaatgcccg 60
tgccagcctc ctggccccac cttcctagtc ctttcctgct gatgacgcta ctgctgggga 120
gactcacagc caaaccctct gcccccgtgg tatcgggccc tgcggtgagg gccacacctg 180
agcacacagt gagcttcacc tgcgagtccc atggcttctc tcccagagac atcaccctga 240
aatggttcaa aaatgggaat gagctctcag acttccagac caacgtggac cccgcaggag 300
acagtgtgtc ctacagcatc cacagcacag ccagggtggt gctgacccgt ggggacgttc 360
actctcaagt catctgcgag atggcccaca tcaccttgca gggggaccct cttcgtggga 420
ctgccaactt gtctgaggcc atccgagttc cacccacctt ggaggttact caacagccca 480
tgagggcaga gaaccaggca aacgtcacct gccaggtgag caatttctac ccccggggac 540
tacagctgac ctggttggag aatggaaatg tgtcccggac agaaacagct tcgaccctca 600
tagagaacaa ggatggcacc tacaactgga tgagctggct cctggtgaac acctgtgccc 660
acagggacga tgtggtgctc acctgtcagg tggagcatga tgggcagcaa gcagtcagca 720
aaagctatgc cctggagatc tcagcacacc agaaggagca cggctcagat atcacccatg 780
aaccagcgct ggctcctact gctccactcc tcgtagctct cctcctgggc cccaagctgc 840
tactggtggt tggtgtctct gccatctaca tctgctggaa acagaaggcc tgactgaccc 900
tcagtctctg ctgcctcctc ctttcttgag aagctcagcc tgagagaagg agctggcgag 960
aaccttcccc acactcagct ccaaacgcct cctctcccag gtcatctgcc tgcccacacg 1020
ctcctgttcc accttcacaa gaccatgatg ccccaaagca gtgtctctat tc 1072

<210> 47
<211> 4822
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505866CB1

<400> 47
aagaaccgac taaggctgtg agtgtccggg aaccagaccc gcttggaggc cacagccccg 60
acgtcccgcg cccacgcggc agatcgggcg ctgcggcctg ggagcctcgg ggagatgcgg 120
agcaggagca attccggggt ccgcttggac gggtacgcgc ggctggtgca gcaaaccatc 180
ctgtgttacc agaatcccgt cacggggctg ctgtcagcca gccatgagca gaaggatgcc 240
tgggtgcggg ataacatcta cagtatcctg gccgtgtggg gcctgggcat ggcctaccgt 300
aagaatgcag accgcgatga ggacaaggcc aaggcctacg agctggagca gaacgtggtg 360
aagctgatgc gaggtcttct ccagtgcatg atgagacagg tggccaaagt ggagaagttc 420
aaacacactc agagcaccaa ggacagcctg cacgccaagt acaacaccgc cacctgtggc 480
acggtggtgg gcgacgacca gtggggccac ctccaggtgg atgccacctc tctcttcctc 540
ctgttcctgg cccagatgac cgcctcaggc ttacgtatca ttttcactct cgatgaggtg 600
gccttcatac agaatcttgt cttttacata gaagctgcat ataaagtcgc tgattatgga 660
atgtgggagc gtggagataa gactaatcag ggcatcccgg aattgaatgc aagctccgta 720
ggaatggcca aggcagctct tgaggcaatt gatgaactgg accttttttgg agcccatgga 780
ggacgcaagt cagtgattca tgttctgcca gatgaggtcg agcactgcca gtctattctg 840
ttctccatgc tgccaagagc gtcgacatct aaagaaattg atgctggact tctttccatt 900
atttccttcc cggcctttgc agtggaagat gtaaaccttg taaatgtgac caaaaatgaa 960
attatttcta agctccaggg gcgttatgga tgctgtcgct tccttcgaga tggttataaa 1020
actccaagag aggaccctaa tcgactgcat tatgaccctg ctgaactcaa gctcttcgaa 1080
aacattgaat gtgagtggcc tgtgttttgg acatatttta taatagatgg agtcttcagt 1140
ggtgatgctg ttcaggtcca agaataccga gaggccctgg agggaatact catcagaggc 1200
aagaatggga tccgcctggt gcctgaactc tacgctgtcc cgcctaacaa ggtagatgaa 1260
gagtacaaga atcctcacac agtagaccga gttcctatgg ggaaggtgcc tcatctgtgg 1320
ggccaatcct tgtacatcct cagctcgctg ttggcagagg gattccttgc cgctggtgaa 1380
atcgatccct taaatagaag attttccact tcagtcaaac ctgatgttgt agtacaagtt 1440
actgttttgg cagaaaacaa tcacattaag gacttattga ggaaacacgg ggtgaacgtc 1500
cagagtatcg cggacattca tccaattcaa gtccagccgg gccggattct tagtcacata 1560
tatgccaagc ttggacggaa taagaatatg aatttgagtg ggcgaccgta tcgacatatt 1620
ggtgtccttg gaacctctaa actatatgtg attaggaacc aaatctttac ttttacaccc 1680
cagttcaccg accagcatca cttctacctg gccctcgaca atgagatgat cgtggagatg 1740
ctaaggatcg agctggccta cctgtgcacc tgctggagga tgacgggcag acccacactc 1800
accttcccca tcagtcgcac catgctcaca aatgatggct cagacattca ttctgctgtg 1860
ctctccacaa ttagaaaact agaggatgga tattttggag gagccagagt aaaattaggg 1920
aacctttcgg aatttctcac cacatcgttc tacacatatc tgacttttct ggatccagac 1980
tgtgatgaga agttgtttga caatgccagc gaagggactt tcagtcctga tagtgattca 2040
gatttggtag gatatctgga agacacctgt aatcaagaaa gccaagacga acttgaccat 2100
tatatcaacc accttctgca aagcacatcg ttgaggtcct atctgcctcc tctttgtaag 2160
aacacagaag accgccatgt cttcagtgct atccactcca cgcgggacat actttctgtg 2220
atggcaaaag caaagggttt ggaagttcca tttgttccca tgactttgcc gactaaagtt 2280
ctaagtgccc accgtaaatc actgaatctt gttgattctc ctcagccact cctagaaaag 2340
gttcctgaaa gtgactttca gtggcccaga gatgaccatg gtgacgtgga ctgtgagaag 2400
ctggttgagc agctaaaaga ttgttcgaac ctacaggacc aagcagacat tctgtacatt 2460
ctttatgtca taaagggtcc cagctgggac acaaatctct ctggacagca cggggtcacc 2520
gttcaaaacc ttcttggtga gctctatggg aaagccggct tgaaccagga gtggggtctg 2580
attcgctaca tctcaggcct tctcaggaag aaagtggagg tcctggctga ggcctgcaca 2640
gacctgcttt cgcaccagaa gcagctcacc gtgggcctgc cgcccgagcc ccgggagaag 2700
atcatctctg cgccccttcc cccagaggag ctcacaaaac tcatctacga ggccagtggg 2760
caggacatca gcattgccgt cctcacgcag gagattgtgg tttacctggc catgtatgtc 2820
agggcgcagc ccagcctctt tgtggagatg ctgagactcc ggattggact gatcattcag 2880
gtgatggcca cggagctggc acggagcctg aactgctcag gagaagaggc ttctgaaagt 2940
ttgatgaacc tcagcccttt cgatatgaaa aatctcctgc accatattct aagtgggaaa 3000
gagtttggcg ttgaaagaag tgtgcgccct atccactcct ccacatccag ccctaccatc 3060
tccatccacg aggtgggcca taccggagtc accaaaactg agaggagtgg cattaacaga 3120
ctgaggagtg aaatgaaaca gaggtccagc accccatcct cgcccactgg cacgtcatcc 3180
tcagactcgg gagggcatca catcggctgg ggtgagcggc agggccagtg gctgcgcagg 3240
agaaggctgg atggggccat caacagggtc cccgtgggat tctaccagag ggtgtggaag 3300

atcctccaga agtgccacgg tctctccatc gatggttatg tcctcccatc ctcgacgacc 3360
cgagagatga ccccgcatga gatcaagttt gctgtccatg tcgaatcggt gctgaaccgc 3420
gtgccgcagc ccgagtaccg gcagctgctg gtggaagcca tcatggtgct gacgctgctc 3480
tcggacacgg agatgaccag catcgggggc atcatccacg tggaccagat cgtgcagatg 3540
gccagtcagc tgttcttgca ggaccaggtg tcaattggtg ccatggacac cctggagaaa 3600
gaccaagcca caggaatctg ccacttcttt tatgacagcg ctccgagtgg ggcttatggg 3660
acgatgacct acctaacaag agcagtggct tcttatttgc aggaattgtt gcccaattcg 3720
ggctgccaga tgcaataggg tctcacctgg aaacatgatc acactctcaa tctgtcacgt 3780
gccccctagc cttactggga accttctgtc ccccaagatc ccctgtgcta tcaggaaagc 3840
atgtcccatc agaaacactc tcgggggca atggtagcac tcaccctgaa actgatgtat 3900
gttaaagcca cagagataga gctgaggagt cctgtgttcc cccgcaagga gcaccccggg 3960
atcattttct aggttcattt ctctggaaca tttgctgtag catctggtct cacggactct 4020
gaggaggaat tggaaattgg tctcttttga gtgcagaggg aactgagacg ccagcttaaa 4080
ttggctcttg cagagagtta cagaaatagt ttcgatgagc tagtgacaca tcctaaagat 4140
gcaaagatcc tcctggcggc agtagccttg acaagggcca cctcttcaca ggatgcagtc 4200
tgtctgtgca ccaaactctt caccaaatag aacacttgtg tctctctgtg gaatgggggt 4260
tttcttgtgc cttgcttgct ttcatagcct tccattttat tggcatggct gccttgatgt 4320
aacataattc tctgtcccca agatttagaa aattcctctt cgttcacctt ggctcatggt 4380
cttccagggt ttttatcctg gctgtctatg aactagggtt ttctcctgcc ttaggaaaaa 4440
tactgcatct tctggaatct agaaaaaaaa aaaaaaaaag acaaggctcc tcttatagcc 4500
aggcaagcag ttgtttaggg ctagacctgt tgtccccgct cgcgtggggt gaacgccccg 4560
cagtggatac tgtcctcctg tgctctgggc gaggcgtcat agggtgagca caagtcagca 4620
gtgccatgga agtgaacagg catgtccccg cagaaacaat gcttccctct gaaaaataga 4680
gccaggctag gaacctgtac cctcttggta gaagcacttt ggtcttcact ctcttgggtg 4740
ggtgcagttg gaggcacgat tggaggagag caaggttgag tcaggtgtca gggtctggaa 4800
atcttttgca atgcccagca ag 4822

<210> 48
<211> 4368
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503214CB1

<400> 48
atgaagtggg taggggacac tggagtgggg ggaaacatcc ctccatcctt cactacccca 60
gggctctcct ccagaccggg tgctatggtg gcggatcgca gccgctggcc actcgcccag 120
gggaagggcg cgcaggcggg cacatggaga gcggcggtgg aatgctccgg ccggggcctc 180
ggggcggcga gcgagtcccc tcagtgcccg ccgccgccgg gggtggaggg cgcggccggg 240
ccggcggagc ccgacggggc ggcggagggc gcggcaggcg gcagcggcga gggcgagagt 300
gggggcgggc cgcggcgggc tctgcgggca gtatacgtgc gcagtgagag ctcccagggc 360
ggcgcggccg gcggcccgga ggctggggcg cggcagtgcc tgctgcgggc ctgcgaggcc 420
gagggcgctc acctcacctc cgtgcccttc ggggagctgg acttcgggga gacggccgtg 480
ctcgacgcct tctacgacgc agatgttgct gtggtagaca tgagcgatgt ctccagacag 540
ccttccctct tctaccatct tggagtccga gaaagctttg acatggccaa taatgtgatc 600
ttgtaccatg acaccgatgc cgacactgct ctctctttga aggacatggt aactcaaaaa 660
aacacagcat ccagtggaaa ttattatttc atcccataca tcgtgacacc gtgcactgat 720
tatttttgct gcgagagtga tgcccagaga cgagcctccg agtacatgca gcccaactgg 780
gacaacatcc tgggcccgct gtgcatgcct ttggtggaca ggttcattag cctccttaag 840
gacatccacg tgacctcatg tgtttattac aaagaaacct tgttaaatga catccggaaa 900
gccagagaga aataccaagg tgaggaactg gcgaaggagc tagctcggat caagctccgc 960
atggataata ctgaggttct gacctcagac atcatcatta acttactcct gtcctaccgt 1020
gatatccagg actatgatgc gatggtgaag ctggtggaaa cactggagat gctgcctacg 1080
tgtgatttgg ccgatcagca taacattaaa ttccactatg cgtttgcact gaataggaga 1140

```
aacagcacag gtgaccgtga gaaggctctg cagatcatgc tccaggttct gcagagctgt 1200
gatcacccgg gccccgacat gttctgcctg tgtgggagga tctacaagga catcttcttg 1260
gattcagact gcaaagatga caccagccgc gacagcgcca ttgagtggta tcgcaaaggg 1320
tttgaactcc agtcatccct ctattcggga attaatcttg cagttttgct gattgttgct 1380
ggacaacaat ttgaaacttc cttggaacta aggaaaatag gtgtccggct gaacagtttg 1440
ttgggaagaa aagggagctt ggagaaaatg aacaattact gggatgtggg tcagttcttc 1500
agcgtcagca tgctggccca tgatgtcggg aaagccgtcc aggcagcaga gaggttgttc 1560
aaactgaaac ctccagtctg gtacctgcga tcattagttc agaacttgtt actaattcgg 1620
cgcttcaaga aaaccattat tgaacactcg cccaggcaag agcggctgaa cttctggtta 1680
gatataattt ttgaggcaac aaatgaagtc actaatggac tcagatttcc agttctggtc 1740
atagagccaa ccaaagtata ccagccttct tatgtttcca taaacaatga agccgaggag 1800
agaacagttt ctttatggca tgtctcaccc acagaaatga aacagatgca cgaatggaat 1860
tttacagcct cttccataaa gggaataaga tttttctctt tggtcaaaga gatgataacc 1920
aatacagcag gcagtacggt ggagctggag ggagagaccg atggagacac cttggagtat 1980
gagtatgacc atgatgcaaa tggtgagaga gttgtcttgg ggaaaggcac gtatgggatt 2040
gtgtatgctg gccgagatct gagcaatcaa gtgcgaatag ccatcaaaga aatcccggag 2100
agagatagca ggtattctca gcctctgcac gaggagatag ccctgcacaa gtaccttaag 2160
caccgcaata tcgttcagta cctgggctct gtttcagaga acggctacat taagatattt 2220
atggagcagg tgcctggagg aagcctttct gctcttctgc gatccaaatg gggcccgatg 2280
aaggaaccga caatcaagtt ttacaccaaa cagatcctgg agggccttaa gtatcttcat 2340
gaaaaccaga tcgtgcacag agacataaag ggcgataatg ttctggtgaa cacctacagc 2400
ggagtggtga aaatctccga ttttggaacc tcgaaacgtc ttgcgggtgt gaacccctgc 2460
acagagactt ttactggcac cctgcagtac atggcacctg agataattga ccaagggcct 2520
cgcggatatg gtgccccagc cgatatctgg tccctgggct gcaccatcat tgagatggcc 2580
accagcaagc ctccgttcca tgagcttggt gagccgcagg cagccatgtt caaagtgggc 2640
atgtttaaga tccaccctga gattccagaa gccctttcag ctgaagcccg agccttcatt 2700
ttatcctgtt tcgagcctga cccccacaaa cgtgccacca ctgctgagct actgagagag 2760
ggtttcttaa ggcaggtgaa caagggcaag aagaaccgaa ttgccttcaa gccctcagaa 2820
ggtccccgcg gtgtcgtcct ggccctgccc acacagggag agcccatggc caccagcagc 2880
agcgagcacg gctctgtctc cccagactcc gacgcccagc ctgacgcact ctttgagagg 2940
acccgggcgc ccaggcacca ccttggccac ctcctcagtg ttccagacga gagctcagcc 3000
ttggaagacc ggggcttggc ctcgtccccg gaggacaggg accagggcct cttcctgcta 3060
cgcaaggaca gtgagcgccg tgccatcctg tacaaaatcc tctgggagga gcagaaccag 3120
gtggcttcca acctgcagga gtgtgtggcc cagagttccg aagagttgca tctctcagtt 3180
ggacacatca agcaaatcat tgggatcctg agggacttca tccgctcccc agagcaccgg 3240
gtgatggcga ccacaatatc aaagctcaag gtggacctgg actttgacag ctcgtccatc 3300
agtcagattc acctggtgct gttcggattt caggatgccg taaataaaat tttgaggaac 3360
cacttaatta ggccccactg gatgttcgcg atggacaaca tcatccgccg agcggtgcag 3420
gccgcggtca ccattctcat cccagagctc cgagcccact ttgagcctac ctgtgagact 3480
gaaggggtag ataaggacat ggatgaagcg gaagagggct atcccccagc caccggacct 3540
ggccaggagg cccagcccca ccagcagcac ctgagcctcc agctgggtga gctcagacag 3600
gagaccaaca gacttttgga acacctagtt gaaaaagaga gagagtacca gaatcttctg 3660
cggcaaactc tagaacagaa aactcaagaa ttgtatcacc ttcagttaaa attaaaatcg 3720
aattgtatta cagagaaccc agcaggcccc tacgggcaga gaacagataa agagcttata 3780
gactggttgc ggctgcaagg agctgatgca aagacaattg aaaagattgt tgaagagggt 3840
tatacacttt cggatattct taatgagatc actaaggaag atctaagata ccttcgacta 3900
cggggtggtc tcctctgcag actctggagt gcggtctccc agtacagaag ggctcaggag 3960
gcctcagaaa ccaaagacaa ggcttgatac caatcagcta agctgtggca gagtgtccca 4020
ccacgctaca tgttttgtta aagcttctgt tagtgtatac acgaattccg ctgtgtttac 4080
atatttaaaa atgccattgt tcaattaata gtttaagaac ttgttttaaa tactgtcctg 4140
agtttctttt gaaacctgtt atttataaac atagaactgt gtgtattgtg aaaacagtga 4200
gccttggttt tgacctcccg gaatattagg aaattcactt gtagtcccag ctatgcagga 4260
ggctgaggtg ggaggattgc ttgagcccag gaggtgtgga ggctgcagtg agccatgatc 4320
acaccactgc actccagcct gggcaacaga gcccgacctg tctcaaaa              4368
```

<210> 49

<211> 2024
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7495312CB1

<400> 49
ataacttgaa gaactgccct ggagaaagga agaaacttat aataaatggg aaattataaa 60
tctagaccaa cccaaacttg tactgatgaa tggaagaaaa aagtcagtga atcatatgtt 120
atcacaatag aaagattaga agatgacctg cagatcaagg aaaaagaact gacagaacta 180
aggaatatat ttggctctga tgaagccttc agtaaagtca atttaaatta ccgcactgaa 240
aatgggctgt ctctacttca tttatgttgc atttgtggag gcaagaaatc acatattcga 300
actcttatgt tgaaagggct ccgcccatct cgactgacaa gaaatggatt tacagccttg 360
catttagcag tttacaagga taatgcagaa ttgatcactt ctctgcttca cagtggagct 420
gatatacagc aggttggata cggtggcctc actgccctcc atattgctac aatagctggc 480
cacctagagg ctgctgatgt gctgttgcaa catggagcta atgtcaatat tcaagatgca 540
gtttttttca ctccattgca tattgcagcg tactatggac atgaacaggt aactcgcctt 600
cttttgaaat ttggtgctga tgtaaatgta agtggtgaag ttggagatag acccctccac 660
ctagcatctg caaaaggatt cttgaatatt gcaaaactct tgatggaaga aggcagcaaa 720
gcagatgtga atgctcaaga taatgaagac catgtcccac tccatttctg ttctcgattt 780
ggacaccatg atatagttaa gtatctgctg caaagtgatt tggaagttca acctcatgtt 840
gttaatatct atggagatac cccccttacac ctggcatgct acaatggcaa atttgaagtt 900
gccaaggaaa tcatccaaat atcaggaaca gaaagtctga ctaaggaaaa catcttcagt 960
gaaacagctt ttcatagtgc ttgtacctat ggcaagagca ttgacctagt caaatttctt 1020
cttgatcaga atgtcataaa catcaaccac caaggaaggg atgggcacac tggattacac 1080
tctgcttgct accacggtca cattcgcctg gttcagttct tactggataa tggagctgat 1140
atgaatctag tggcttgtga tcccagcagg tctagtggtg aaaaagatga gcagacatgt 1200
ttgatgtggg cttatgaaaa agggcatgat gccattgtca cactcctgaa gcattataag 1260
agaccacaag atgaattgcc ctgtaatgaa tattctcagc ctggaggaga tggctcctat 1320
gtgtctgttc catcaccctt ggggaagatt aaaagcatga caaaagagaa ggcagatatt 1380
ctcctcctaa gagctggatt gccttcacat ttccatcttc agctctcaga aattgagttc 1440
catgagatta ttggctcagg ttcttttggg aaagtatata aaggacgatg cagaaataaa 1500
atagtggcta taaaacgtta tcgagccaat acctactgct ccaagtcaga tgtggatatg 1560
ttttgccgag aggtgtccat tctctgccag ctcaatcatc cctgcgtaat tcagtttgtg 1620
ggtgcttgct tgaatgatcc cagccagttt gccattgtca ctcaatacat atcaggggt 1680
tctctgttct ccctccttca tgagcagaag aggattcttg atttgcagtc taaattaatt 1740
attgcagtag atgttgccaa aggcatggag taccttcaca acctgacaca gccaattata 1800
catcgtgact tgaacagatg ctgtacagga tggctcagct gctaccaccc agactgataa 1860
actgggtcat ctggtctcct ggccccagga gactgactca gcatcaagag gacaactttg 1920
acaccctatg atatcatctc caacccaacc aatcagcagt ccccactttc agacccctac 1980
ccaccaagtt ttccttaaaa cccccatcct caagtttttg agaa 2024

<210> 50
<211> 715
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506732CB1

<400> 50
cggatgctgc ggcgcccgct ggccgggctg gctgcggccg ccctgggccg ggccccaccg 60
gacggcttgc tctgctcttt acctggggtt gctgtcgagg accctgtgca agactcggcc 120

ggtttttctt tctccctgat ggacagaccc aaacatagcc gcgcagcatc gtgaagggct 180
ggggccttca ctcctctgtg gctctggaag agcccgattt cctcaggagg catgtcgggc 240
cccaggcctg tggtgctgag cgggccttcg ggagctggga agagcaccct gctgaagagg 300
ctgctccagg agcacagcgg catctttggc ttcagcgtgt cccataccac gaggaacctg 360
tatggcacga gcaaggtggc ggtgcaggcc gtgcaggcca tgaaccgcat ctgtgtgctg 420
gacgtggacc tgcagggtgt gcggaacatc aaggccaccg atctgcggcc catctacatc 480
tctgtgcagc cgccttcact gcacgtgctg gagcagcggc tgcggcagcg caacactgaa 540
accgaggaga gcctggtgaa gcggctggct gctgcccagg ccgacatgga gagcagcaag 600
gagcccggcc tgtttgatgt ggtcatcatt aacgacagcc tggaccaggc ctacgcagag 660
ctgaaggagg cgctctctga ggaaatcaag aaagctcaaa ggaccggcgg ctgag 715

<210> 51
<211> 591
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506736CB1

<400> 51
cggatgctgc ggcgcccgct ggccgggctg gctgcggccg ccctgggccg ggccccaccg 60
gacggcttgc tctgctcttt acctggggtt gctgtcgagg accctgtgca agactcggcc 120
ggtttttctt tctccctgat ggacagaccc aaacatagcc gcgcagcatc gtgaagggct 180
ggggccttca ctcctctgtg gctctggaag agcccgattt cctcaggagg catgtcgggc 240
cccaggcctg tggtgctgag cgggccttcg ggagctggga agagcaccct gctgaagagg 300
ctgctccagg agcacagcgg catctttggc ttcagcgtgt cccataccac gaggaacccg 360
aggcccggcg aggagaacgg caaagattac tactttgtaa ccagggaggt gatgcagcgt 420
gacatagcag ccggcgactt catcgagcat gccgagttct cgaggagagc agcaaggagc 480
ccggcctgtt tgatgtggtc atcattaacg acagcctgga ccaggcctac gcagagctga 540
aggaggcgct ctctgaggaa atcaagaaag ctcaaaggac cggcggctga g 591

<210> 52
<211> 7424
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507121CB1

<400> 52
cggtgccggg gtgggggtga ggctcagcag agcccggggc acaggccgga ggggctgcag 60
gcgcgaggct gggttacgtg aggaagctgg gggtttcgcg ggcagcttta gagccccagt 120
cagggaaacc gaggccgggc ttcctggctg cctcgcgagc ctcttcatgg ctctcgccgc 180
cgccctgagg tgcctagaat gggttccggc ctccggggag gttcccagta accgcaggag 240
ccaccattga tttggcgtct gctgggtgca aagcccagcg cgctaaccct ttactcgcga 300
cctttcgctt caccttcaca gcagccctgc gaggagagtt gtggactggg gcaacctttg 360
ccagtgatga gaagtgatgc tcgtggcagt gctgaatctc tctgaatatg attcgaattg 420
cagccttaaa tgccagctcc accattgagg atgatcatga aggaagcttt aaaagtcaca 480
aaacccagac aaaggaggct caggaagcag aggcttttgc attgtaccac aaggcccttg 540
atctgcagaa acatgaccgg tttgaggagt ctgccaaagc ctaccatgag ctcttggagg 600
cgagcctgct gcgggaggca gtttcatccg gtgatgagaa agaggggttg aaacaccctg 660
ggctgatact gaaatattcc acttataaga acttggccca gctggcagcc cagcgggagg 720
atctggagac agccatggag ttctacttag aggcagtgat gctggactcc acagatgtca 780
acctctggta taagattgga catgtggccc tgaggctcat ccggatcccc ctggctcgcc 840

```
atgcttttga ggaagggctg cggtgcaatc ctgaccactg gccctgtttg gataacctaa 900
tcactgtcct gtacaccctc agtgattaca caacatgtct gtacttcatc tgcaaagctt 960
tggagaagga ttgccggtac agcaaagggc tggtcctcaa ggagaagatt tttgaggagc 1020
agccttgtct ccggaaggac tctctcagaa tgttcctcaa atgtgacatg tcgattcacg 1080
atgtttcggt gagtgcagct gagacacagg cgattgtaga tgaggccttg ggctgcgaa 1140
aaaagaggca agcgctgatt gtgcgggaga aggagccgga cctgaaactt gtgcagccca 1200
ttcctttctt cacctggaag tgcctcggag agagcttgct ggccatgtac aatcatctca 1260
ccacctgtga gcccccacgt cccagccttg gcaaaaggat tgatttgtcg gactaccagg 1320
accccagcca gcctcttgag tcctccatgg tggtgacgcc agttaacgtg atccagccaa 1380
gcactgtcag caccaaccca gctgtggctg tcgccgagcc tgtggtctcc tacacctctg 1440
tggctacaac cagcttccca ctgcacagtc ctggtctgtt ggagacaggc gctcctgtgg 1500
gtgatatttc tgggggagat aaatccaaga aggggtaaa acggaagaag atttcagaag 1560
agagtggaga aacagcaaag cggcggtctg cccgtgtccg aaacaccaag tgcaaaaaag 1620
aagagaaagt agacttccag gagcttctga tgaagttctt gccgtccagg ttaagaaagc 1680
tggaccctga ggaggaagat gattccttta ataactatga agtccagtca gaagccaaac 1740
tggaaagctt cccaagcatt gggcctcaaa ggctgtcatt tgactcagcc acattcatgg 1800
aatctgaaaa gcaggacgtg catgagttcc tgctggagaa cctaaccaac gggggcatcc 1860
tggagctgat gatgcgctac ctgaaagcca tgggccacaa gttcttggta aggtggcctc 1920
caggcttggc ggaggtcgtg ctcagcgtct accacagctg gaggaggcac agcaccagcc 1980
tgcccaaccc gctgctgagg gactgcagca acaagcacat caaggacagg atgctgatgt 2040
ctctctcctg catggaactc cagctggacc agtggctgct gaccaaaggc agaagctctg 2100
cagtgtctcc tcggaactgc cctgctggta tggtgaatgg cagatttgga cctgacttcc 2160
cagggaccca ctgcctgggt gacctcctac agctgtcatt tgcctcgtcc cagcgcgacc 2220
tgttcgagga tggttggctg gagtttgtgg tccgtgttta ctggctgaag gctcgcttcc 2280
tggcgctgca gggagacatg gagcaggccc tggagaacta tgacatctgc acagaaatgc 2340
tccagagttc caccgccatc caggtggagg caggggctga acgaagagac attgtcatcc 2400
ggctgcccaa cctccataat gactctgtgg tttccctgga ggagattgat aagaacctga 2460
agtcgctgga gcggtgccag tccctggagg agattcagcg gctgtatgaa gcaggcgact 2520
acaaggctgt tgtgcatctg ctccgcccca ctttgtgcac cagtgggttt gaccgggcca 2580
aacacctgga gtttatgact tccattcctg agaggccagc ccagctgctt cttctgcagg 2640
actccttgct ccggctgaag gactatcggc agtgttttga gtgttccgat gtggctctga 2700
acgaggctgt ccagcagatg gtgaactcag gtgaggctgc cgccaaggag gagtgggtgg 2760
ccacagtgac ccaactgctg atgggcatcg agcaggccct ctctgcggac agcagtggta 2820
gcatcctgaa ggtatcatcc tccaccactg gccttgtgcg gctcaccaac aacctcatcc 2880
aggtcattga ctgcagcatg gctgtgcagg aggaggccaa ggagccccac gtctcttcag 2940
tgctaccctg gatcattcta caccggatca tctggcagga ggaagacacc ttccattctc 3000
tgtgccacca gcagcagctc caaaacccag cggaggaagg gatgtcagag acgcccatgc 3060
tcccatcctc cctcatgctg ctgaacacag cccacgagta tttgggcaga aggtcctggt 3120
gctgcaattc agatggggct ctgctgcgat tctatgtgcg agtactccag aaggaactgg 3180
ctgcatccac ctctgaagac acgcaccctt acaaggagga gctggagaca gccttggagc 3240
agtgcttcta ctgcctgtac agcttcccca gcaagaagag taaggccagg tacctggagg 3300
aacactcggc ccagcaggtg gatcttatat gggaggatgc actgttcatg tttgagtatt 3360
ttaagcccaa gacccttcct gaatttgaca gctataagac cagcaccgtg tctgctgact 3420
tggccaacct actgaagaga attgccacca ttgtgcctcg cacagagagg ccagccctta 3480
gcctggacaa agtctctgcc tacattgagg gaacttcaac tgaggtaccc tgcctcccag 3540
agggggctga cccctcccct ccagtggtga acgagcttta ctacctcctg gctgattatc 3600
atttcaaaaa caaggagcag tccaaggcca tcaagttcta catgcatgac atctgcatct 3660
gccccaatag gtttgattcc tgggcaggca tggctctggc ccgggccagc cgcattcagg 3720
acaagctgaa ctccaatgag ctgaagagtg atgggcccat ttggaagcat gccacgcccg 3780
tcttgaactg cttccgtcgg gccctggaga ttgacagctc caacttgtcc ctatggattg 3840
agtatggcac catgtcctat gccttgcact cattcgcctc acgtcaattg aagcagtgga 3900
gaggcgagct gccccctgag ctcgtgcagc agatggaggg ccggcgcgac agcatgctag 3960
agacagccaa gcactgtttc acatcagcag cccgctgcga gggtgatggt gacgaggagg 4020
agtggctcat ccactacatg ctgggcaagg tggctgagaa gcagcagcag ccacccaccg 4080
tttacttgct gcactacagg caggctggcc actacctgca cgaggaggct gcccgctacc 4140
ccaagaagat ccactaccac aacccacctg agctggccat ggaggccctg gaggtgtact 4200
```

ttcggctcca tgcttccatc ctgaagctcc tggggaagcc cgattctggg gttggtgcag 4260
aggtcctggt caactttatg aaggaggctg cagaaggacc ctttgccagg ggcgaggaga 4320
agaacacacc caaagcttca gaaaaggaga aggcctgcct ggtggacgag gactcccact 4380
cttcagctgg gacactgccg ggccccggag cctccctccc ctcctcctct ggcccaggtc 4440
tgacatcccc accttacaca gccactccga ttgaccacga ttacgtcaaa tgtaaaaaac 4500
cccaccagca ggcaacgccg gacgaccgaa gccaggacag cacagccgta gcactctcag 4560
actctagctc aacgcaggac ttctttaatg agcccaccag cttactggaa ggctccagga 4620
aatcctacac agagaagagg ctgcccattc tcagttccca agcaggagcg acgggtaaag 4680
atcttcaggg ggccacagaa gaaagaggaa aaaacgagga gtcattggag agtacagaag 4740
gcttccgggc tgcagagcaa ggtgtccaga agcctgctgc agaaacccca gcctctgctt 4800
gcatccctgg caagccctca gcatccacac ccaccctgtg ggatgggaag aagagagggg 4860
acctcccagg ggagccagtg gccttccccc aggggctgcc ggctggtgct gaggagcagc 4920
ggcagtttct cacagagcag tgcatcgcct ccttccgcct gtgcctgagc cgcttccccc 4980
agcactataa gagtctctac cgtctggcct tcctctacac ctacagcaag acccaccgga 5040
acctccagtg ggcccgcgac gtgttgctag gcagcagtat cccgtggcaa caactgcagc 5100
acatgccggc acaggggctc ttctgcgaga ggaacaagac caatttcttc aacggcatct 5160
ggcggatccc cgtggacgag attgaccggc cgggcagctt tgcctggcac atgaaccgct 5220
ccatcgtgct gctgctcaag gtgctggccc agctgcggga ccacagcacc ctgctgaagg 5280
tgtcctccat gcttcagcgg accccagacc agggcaagaa gtatctgcga gatgctgacc 5340
gccaggtcct ggcgcagcgg gccttcatcc tcactgtgaa ggtgctcgaa gacacgctga 5400
gcgagctcgc agaggggtca gaacgcccag ggcccaaggt ctgtggcctc cccggagcca 5460
ggatgaccac cgatgtctca cacaaggcca gtcctgagga tggccaggag ggcctccccc 5520
agccgaagaa gccccctctg gctgatggct cagggccagg gcccgagcca ggaggcaaag 5580
tgggcctcct caaccaccgg cctgtggcca tggatgcagg agacagtgca gaccaaagcg 5640
gggagcggaa ggataaagag agcccacggg cagggcccac tgagcccatg gacacgagtg 5700
aggccactgt ttgccactca gacttggagc ggacaccacc cctgctgcca ggtcgccccg 5760
caagggaccg gggccccgag agccggccca ctgagctgtc cctggaggag ctgagcatca 5820
gtgcccggca gcagcccacc ccgctcaccc cagcccagcc agcccccgcc cccgccccg 5880
ccaccaccac agggaccagg gcagggggcc acccggagga gccgctctcc cggctcagcc 5940
gcaagaggaa gctcctggag gacacagagt caggcaagac acttctgttg gatgcctacc 6000
gtgtgtggca gcagggccag aagggtgtgg cctatgacct gggccgtgtg gagaggatca 6060
tgtcggagac ctacatgctc atcaagcagg tggatgagga ggctgcgctg gagcaggctg 6120
tgaagttctg ccaggtccat cttggggctg ccgcccagag acaggcctcg ggggacaccc 6180
ccaccactcc aaagcacccc aaagacagcc gagagaactt ctttcctgtg acagtggtgc 6240
ccacagcccc tgaccctgtg ccagctgact ctgtccagcg gcccagtgat gctcacacca 6300
agcctcgccc tgcactagct gccgccacaa ctattatcac ctgccctccg tcagcatcag 6360
cttccaccct ggaccagtcc aaggaccctg ggcctccccg gccacacagg cctgaagcta 6420
cccccagcat ggcctctctg ggcccagagg gagaagagct ggcgagagtg gcagagggca 6480
ccagcttccc gcctcaggag ccacggcaca gtccgcaggt gaagatggcc cccacaagtt 6540
ccccggcaga gccacactgc tggccggcag aggctgccct gggcacaggc gctgagccca 6600
cctgcagcca ggctgccagc tccaaggccc ccagcagtgg gagtgcccag ccaccagagg 6660
gtcacccagg caagcctgag cccagccggg ctaagtcccg cccccctgccc aacatgccaa 6720
agctggtcat cccctccgcc gccaccaagt tccccccctga gatcaccgtc acgccaccca 6780
ccccaaccct gctctccccc aaaggcagca tctcggagga gaccaagcag aagctgaagt 6840
cagccatcct ttctgcccag tctgctgcca acgtgaggaa ggagagccta tgccagccag 6900
ccctggaggt cctggagaca tccagccagg agtcctcgct ggagagcgag acagacgagg 6960
acgacgacta catggacatt tgaggggcca ctgcagcccc accgccacgc cccaggggac 7020
cagccaggcc tggaatgccc cctgggcagg accctgggca ggaccagagg cccacatgga 7080
tgccactccc cacacagccc ccaggcctgc ccagcccacc tcctcatggc atcctccctg 7140
tacccaggtc aggctgtcca caccacatgg gagcccagag gaggaggggc ccgccttagc 7200
catgtgaagg tggattggtc gccatctgca cgccaggcgg catccttttc tatgaagtgt 7260
tgactttgta aatctgccca cacccagctg gccatatcca cccctcgacg ccgggatgag 7320
ccggctctgc ctgtgtcaca gtggaggggt cctttagggc caggctcacc cctcaccctt 7380
tttttggttg cttttctaat aaagatggaa cagttaaaaa aaaa 7424

<210> 53

<211> 1303
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 90086258CB1

<400> 53
ggagctttat ttcatttggg atttcaagtt tacagatggt atcttctcaa aagttggaaa 60
aacctataga gatgggcagt agcgaacccc ttcccatcgc agatggtggc aggaggagga 120
agaagaagcg gaggggccgg gccactgact ccttgccagg aaagtttgaa gatatgtaca 180
agctgacctc tgaattgctt ggagagggag cctatgccaa agttcaaggt gccgtgagcc 240
tacagaatgg caaagagtat gccgtcaaaa tcatcgagaa acaagcaggg cacagtcgga 300
gtagggtgtt tcgagaggtg gagacgctgt atcagtgtca gggaaacaag aacattttgg 360
agctgattga gttctttgaa gatgacacaa ggttttactt ggtctttgag aaattgcaag 420
gaggttccat cttagcccac atccagaagc aaaagcactt caatgagcga gaagccagcc 480
gagtggtgcg ggacgttgct gctgcccttg acttcctgca taccaaaggc attgctcatc 540
gtgatctgaa accagaaaat atattgtgtg aatctccaga aaaggtgtct ccagtgaaaa 600
tctgtgactt tgacttgggc agtgggatga aactgaacaa ctcctgtacc cccataacca 660
caccagagct gaccacccca tgtggctctg cagaatacat ggcccctgag gtagtggagg 720
tcttcacgga ccaggccaca ttctacgaca agcgctgtga cctgtggagc ctgggcgtgg 780
tcctctacat catgctgagt ggctacccac ccttcgtggg tcactgcggg gccgactgtg 840
gctgggaccg gggcgaggtc tgcagggtgt gccagaacaa gctgtttgaa agcatccagg 900
aaggcaagta tgagtttcct gacaaggact gggcacacat ctccagtgaa gccaaagacc 960
tcatctccaa gctcctggtg cgagatgcaa agcagagact tagcgccgcc caagttctgc 1020
agcacccatg ggtgcagggg gaacagcagc acaatggacc tgacgctctt cgcagctgag 1080
gccatcgccc ttaaccgcca gctatctcag cacgaagaga acgaactagc agaggagcca 1140
gaggcactag ctgatggcct ctgctccatg aagccttccc ctccctgcaa gtcacgcctg 1200
gcccggagac gggccctggc ccaggcaggc cgtggtgaag acaggagccc gcccacagca 1260
ctctgaaatg ctccagtcac accttattgg ccctaggcct ggc 1303

<210> 54
<211> 2111
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1967990CB1

<400> 54
ggcgtacgtg ctcgctcacg gcggcggcgg cggagcggag aggccagagc cggagaccga 60
gctgggatcg ggccccgggc gggggcggtg cgagcggcgc caagcagatc ttaggggcgg 120
ggacggagcc ggggcgggcg ggactgaagc ggagcccggg aacggggcgg gaggtcccag 180
tgtcccgggt tgggggggtg gagcagcatt tcgtcgccgc gggggtgccg ggactccggc 240
cgcagtgtcg ccgccatcac ggacttcctg tgggacaagc gcacgggcct cgccgccaga 300
acgatgccgc atcctcgaag gtaccactcc tcagagcgag gcagccgggg gagttaccgt 360
gaacactatc ggagccgaaa gcataagcga cgaagaagtc gctcctggtc aagtagtagt 420
gaccggacac gacggcgtcg gcgagaggac agctaccatg tccgttctcg aagttatgat 480
gatcgttcgt ccgaccggag ggtgtatgac cggcgatact gtggcagcta cagacgcaac 540
gattatagcc gggatcgggg agatgcctac tatgacacag actatcggca ttcctatgaa 600
tatcagcggg agaacagcag ttaccgcagc cagcgcagca gccggaggaa gcacagacgg 660
cggaggaggc gcagccggac atttagccgc tcatcttcgc agcacagcag ccggagagcc 720
aagagtgtag aggacgacgc tgagggccac ctcatctacc acgtcgggga ctggctacaa 780
gagcgatatg aaatcgttag caccttagga gaggggacct tcggccgagt tgtacaatgt 840

gttgaccatc gcaggggtgg ggctcgagtt gccctgaaga tcattaagaa tgtggagaag 900
tacaaggaag cagctcgact tgagatcaac gtgctagaga aaatcaatga gaaagaccct 960
gacaacaaga acctctgtgt ccagatgttt gactggtttg actaccatgg ccacatgtgt 1020
atctcctttg agcttctggg ccttagcacc ttcgatttcc tcaaagacaa caactacctg 1080
ccctacccca tccaccaagt gcgccacatg gccttccagc tgtgccaggc tgtcaagttc 1140
ctccatgata acaagctgac acatacagac ctcaagcctg aaaatattct gtttgtgaat 1200
tcagactatg agctcaccta caacctagag aagaagcgag atgagcgcag tgtgaagagc 1260
acagctgtgc gggtggtaga ctttggcagt gccacctttg accatgagca ccatagcacc 1320
attgtctcca ctcgccatta ccgagcacca gaagtcatcc ttgagttggg ctggtcacag 1380
ccttgtgatg tgtggagtat aggctgcatc atctttgaat actatgtggg attcaccctc 1440
ttccagaccc atgacaacag agagcatcta gccatgatgg aaaggatctt gggtcctatc 1500
ccttcccgga tgatccgaaa gacaagaaag cagaaatatt tttaccgggg tcgcctggat 1560
tgggatgaga acacatcagc tgggcgctat gttcgtgaga actgcaaacc gctgcggcgg 1620
tatctgacct cagaggcaga ggaacaccac cagctcttcg atctgattga aagcatgcta 1680
gagtatgaac cagctaagcg gctgaccttg ggtgaagccc ttcagcatcc tttcttcgcc 1740
cgccttcggg ctgagccgcc caacaagttg tgggactcca gtcgggatat cagtcggtga 1800
cgatcaggcc ctgggccccc ctgcatcttt tatagcagtg ggtgtccagt ccaggacact 1860
ggtgcttttt tatacaagag aacgagccag agttcactcc ttcctcctgg ctctctatat 1920
acctgtgaat atgtgaaata gtgtaaatat gaaagaactt gtacctatca cttcaacccc 1980
tgccttgtac ataatactat tccatccaca cagtttccac cctcacctgc cccctcatac 2040
ggagttggat gggggccgag tgaggtaacc aggtggcatc taccccatgt tttataagga 2100
attttgtaca g 2111

<210> 55
<211> 4017
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3810039CB1

<400> 55
cgttgggggg gggtcctccc agctctgcgg aggcgcatga acaataggcg gcggcggctc 60
ctgcgggcgg cggcggcccc gcaccctatg gatcggccgc tccatggagc cctccagagc 120
gcttctcggc tgcctagcga gcgccgccgc tgccgccccg ccgggggagg atggagcagg 180
ggccggggcc gaggaggagg aggaggagga ggaggaggcg gcggcggcgg tgggccccgg 240
ggagctgggc tgcgacgcgc cgctgcccta ctggacggcc gtgttcgagt acgaggcggc 300
gggcgaggac gagctgaccc tgcggctggg cgacgtggtg gaggtgctgt ccaaggactc 360
gcaggtgtcc ggcgacgagg gctggtggac cgggcagctg aaccagcggg tgggcatctt 420
ccccagcaac tacgtgaccc cgcgcagcgc cttctccagc cgctgccagc ccggcggcga 480
ggaccccagt tgctacccgc ccattcagtt gttagaaatt gattttgcgg agctcacctt 540
ggaagagatt attggcatcg gggctttgg gaaggtctat cgtgctttct ggataggga 600
tgaggttgct gtgaaagcag ctcgccacga ccctgatgag gacatcagcc agaccataga 660
gaatgttcgc caagaggcca agctcttcgc catgctgaag caccccaaca tcattgccct 720
aagaggggta tgtctgaagg agcccaacct ctgcttggtc atggagtttg ctcgtggagg 780
acctttgaat agagtgttat ctgggaaaag gattccccca gacatcctgg tgaattgggc 840
tgtgcagatt gccagaggga tgaactactt acatgatgag gcaattgttc ccatcatcca 900
ccgcgacctt aagtccagca acatattgat cctccagaag gtggagaatg gagacctgag 960
caacaagatt ctgaagatca ctgattttgg cctggctcgg gaatggcacc gaaccaccaa 1020
gatgagtgcg gcagggacgt atgcttggat ggcacccgaa gtcatccggg cctccatgtt 1080
ttccaaaggc agtgatgtgt ggagctatgg ggtgctactt tgggagttgc tgactggtga 1140
ggtgcccttt cgaggcattg atggcttagc agtcgcttat ggagtggcca tgaacaaact 1200
cgcccttcct attccttcta cgtgcccaga accttttgcc aaactcatgg aagactgctg 1260
gaatcctgat ccccactcac gaccatcttt cacgaatatc ctggaccagc taaccaccat 1320
agaggagtct ggtttctttg aaatgcccaa ggactccttc cactgcctgc aggacaactg 1380

gaaacacgag attcaggaga tgtttgacca actcagggcc aaagaaaagg aacttcgcac 1440
ctgggaggag gagctgacgc gggctgcact gcagcagaag aaccaggagg aactgctgcg 1500
gcgtcgggag caggagctgg ccgagcggga gattgacatc ctggaacggg agctcaacat 1560
catcatccac cagctgtgcc aggagaagcc ccgggtgaag aaacgcaagg gcaagttcag 1620
gaagagccgg ctgaagctca aggatggcaa ccgcatcagc ctcccttctg atttccagca 1680
caagttcacg gtgcaggcct cccctaccat ggataaaagg aagagtctta tcaacagccg 1740
ctccagtcct cctgcaagcc ccactatcat tcctcgcctt cgagccatcc agttgacacc 1800
aggtgaaagc agcaaaacct ggggcaggag ctcagtcgtc ccaaaggagg aaggggagga 1860
ggaggagaag agggccccaa agaagaaggg acggacgtgg gggccaggga cgcttggtca 1920
gaaggagctt gcctcgggag atgaaggcct caagtccctg gtagatggat ataagcagtg 1980
gtcgtccagt gcccccaacc tggtgaaggg cccaaggagt agcccggccc tgccagggtt 2040
caccagcctt atggagatgg aggatgagga cagtgaaggc ccagggagtg gagagagtcg 2100
cctacagcat tcacccagcc agtcctacct ctgtatccca ttccctcgtg gagaggatgg 2160
cgatggcccc tccagtgatg gaatccatga ggagcccacc ccagtcaact cggccacgag 2220
taccccctcag ctgacgccaa ccaacagcct caagcggggc ggtgcccacc accgccgctg 2280
cgaggtggct ctgctcggct gtggggctgt tctggcagcc acaggcctag ggtttgactt 2340
gctggaagct ggcaagtgcc agctgcttcc cctggaggag cctgagccac cagcccggga 2400
ggagaagaaa agacgggagg gtctttttca gaggtccagc cgtcctcgtc ggagcaccag 2460
ccccccatcc cgaaagcttt tcaagaagga ggagcccatg ctgttgctag gagacccctc 2520
tgcctccctg acgctgctct ccctctcctc catctccgag tgcaactcca cacgctccct 2580
gctgcgctcc gacagcgatg aaattgtcgt gtatgagatg ccagtcagcc cagtcgaggc 2640
ccctcccctg agtccatgta cccacaaccc cctggtcaat gtccgagtag agcgcttcaa 2700
acgagatcct aaccaatctc tgactcccac ccatgtcacc ctcaccaccc cctcgcagcc 2760
cagcagtcac cggcggactc cttctgatgg ggcccttaag ccagagactc tcctagccag 2820
caggagcccc tccagcaatg ggttgagccc cagtcctgga gcaggaatgt tgaaaacccc 2880
cagtcccagc cgagacccag gtgaattccc ccgtctccct gaccccaatg tggtcttccc 2940
cccaacccca aggcgctgga acactcagca ggactctacc ttggagagac ccaagactct 3000
ggagtttctg cctcggccgc gtccttctgc caaccggcaa cggctggacc cttggtggtt 3060
tgtgtccccc agccatgccc gcagcacctc cccagccaac agctccagca cagagacgcc 3120
cagcaacctg gactcctgct ttgctagcag tagcagcact gtagaggagc ggcctggact 3180
tccagccctg ctcccgttcc aggcagggcc gctgcccccg actgagcgga cgctcctgga 3240
cctggatgca gaggggcaga gtcaggacag caccgtgccg ctgtgcagag cggaactgaa 3300
cacacacagg cctgcccctt atgagatcca gcaggagttc tggtcttagc acgaaaagga 3360
ttggggcggg caagggggac agccagcgga gatgagggga gctggcgggc acagcccttt 3420
ctcagggttg gacccctga gatccagccc tacttcttgc actgataatg cactttgaag 3480
atggaaggga tggaaacagg gccacttcag agggtctcct gccctgcagg gcctttctac 3540
ccgtgtccac tggaggggct gtggccatca gctctggctg tgtaggggag gaaggggtgc 3600
atgcatgtcc cccaccctcc acagtcttcc ttgcctttag agtgaccctg cagagtcact 3660
cagccaaatc tgtctgctgc tccctctcct cagccagttg ggtgtgcgca gagctgtcat 3720
agggtccctt tgtcagcccc gagttcagct tcccaaacac cagtgttgga tattctgtga 3780
ttgattttgg tcctcctccg ctgtccccca acacccagga atgggaatct ggcttggttc 3840
gagataggag cttttctgtg tcctaagccc tttcatgcta gcaggaagac tgaaagcaag 3900
gtggcccagt gtggggtcat agggcttgat agacctggca ctgcctatct gcacttccag 3960
gtgccccacc tatttatctg agcccacagg tggaaagggg aactgcctca gtgagaa 4017

<210> 56
<211> 718
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8032337CB1

<400> 56
ccatccgggt actctaagtc ctccctcccc cttctccttc acccccaagt gtctcccacg 60

actgccctcc ccgacctcta gctgaccatg gcatcagagg cagaaaaaac attccatcgg 120
tttgctgcgt ttggagaatc atcaagcagt ggcactgaaa tgaacaacaa gaacttctcc 180
aagctgtgca aagactgtgg catcatggat ggcaagacag tcacctccac ggacgtggac 240
atcgtgttca gcaaagtcaa ggccaagaac gcccgaacca tcacgtttca acagttcaaa 300
gaggcagtga aggaactggg ccagaagcgc ttcaaaggga agagtccaga tgaagtcctg 360
gagaacattt atggactcat ggagggcaaa gacccagcca ccactggcgc tactaaagca 420
acaacagtgg gtgcagtgga ccgtttgaca gacaccagca agtacaccgg cacccacaag 480
gagcgctttg atgagagtgg caagggcaag ggcattgcgg gacgggaaga gatgactgac 540
aacacaggct atgtgagtgg ttacaagggt tctggcacct acgataagaa gaccaagtag 600
agaggagctt catctcagcc tgctagcccc ctgaccctgc atgtttaaca ccagggagct 660
tggaaaacaa taaacatctg tgtgtgcagc aaaaaaaaaa aaaaaaaaaa aaaaattg 718

<210> 57
<211> 4108
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506411CB1

<400> 57
ggcgaggtca cgcctctctt cagacagccc gagccttccc ggcctggcgc gtttagttcg 60
gaactgcggg acgccggtgg gctagggcaa ggtgtgtgcc ctcttcctga ttctggagaa 120
aaatgccggt ccggaagcaa gatacccaga gagcattgca ccttttggag gaatatcgtt 180
caaaactaag ccaaactgaa gacagacagc tcagaagttc catagaacgg gttattaaca 240
tatttcagag caacctcttt caggctttaa tagatattca agaattttat gaagtgacct 300
tactggataa tccaaaatgt atagatcgtt caaagccgtc tgaaccaatt caacctgtga 360
atacttggga gatttccagc cttccaagct ctactgtgac ttcagagaca ctgccaagca 420
gccttagccc tagtgtagag aaatacaggt atcaggatga agatacacct cctcaagagc 480
atatttcccc acaaatcaca aatgaagtga taggtccaga attggttcat gtctcagaga 540
agaacttatc agagattgag aatgtccatg gatttgtttc tcattctcat atttcaccaa 600
taaaggcaaa tcctccccca gtactggtca acacagatag cttggaaaca ccaacttacg 660
ttaatggcac agatgcagat tatgaatatg aagaaatcac acttgaaagg ggaaattcag 720
ggcttggttt cagcattgca ggaggtacgg acaacccaca cattggagat gactcaagta 780
ttttcattac caaaattatc acagggggag cagccgccca agatggaaga ttgcgggtca 840
atgactgtat attacgagta aatgaagtag atgttcgtga tgtaacacat agcaaagcag 900
ttgaagcgtt gaaagaagca gggtctattg tacgcttgta tgtaaaaaga aggaaaccag 960
tgtcagaaaa aataatggaa ataaagctca ttaaaggtcc taaaggtctt gggtttagca 1020
ttgctggagg tgttggaaat cagcatattc ctggggataa tagcatctat gtaaccaaaa 1080
taattgaagg aggtgcagca cataaggatg gcaaacttca gattggagat aaacttttag 1140
cagtgaataa cgtatgttta gaagaagtta ctcatgaaga agcagtaact gccttaaaga 1200
acacatctga ttttgtttat ttgaaagtgg caaaacccac aagtatgtat atgaatgatg 1260
gctatgcacc acctgatatc accaactctt cttctcagcc tgttgataac catgttagcc 1320
catcttcctt cttgggccag acaccagcat ctccagccag atactcccca gtttctaaag 1380
cagtacttgg agatgatgaa attacaaggg aacctagaaa agttgttctt catcgtggct 1440
caacgggcct tggtttcaac attgtaggag gagaagatgg agaaggaata tttatttcct 1500
ttatcttagc cggaggacct gctgatctaa gtggagagct cagaaaagga gatcgtatta 1560
tatcggtaaa cagtgttgac ctcagagctg ctagtcatga gcaggcagca gctgcattga 1620
aaaatgctgg ccaggctgtc acaattgttg cacaatatcg acctgaagaa tacagtcgtt 1680
ttgaagctaa aatacatgat ttacgggagc agatgatgaa tagtagtatt agttcagggt 1740
caggttctct tcgaactagc cagaagcgat ccctctatgt cagagccctt tttgattatg 1800
acaagactaa agacagtggg cttcccagtc agggactgaa cttcaaattt ggagatatcc 1860
tccatgttat taatgcttct gatgatgaat ggtggcaagc caggcaggtt acaccagatg 1920
gtgagagcga tgaggtcgga gtgattccca gtaaacgcag agttgagaag aaagaacgag 1980
cccgattaaa aacagtgaaa ttcaattcta aaacgagaga taaagggcag tcattcaatg 2040

acaagcgtaa aaagaacctc ttttcccgaa aattcccctt ctacaagaac aaggaccaga 2100
gtgagcagga aacaagtgat gctgaccagc atgtaacttc taatgccagc gatagtgaaa 2160
gtagttaccg tggtcaagaa gaatacgtct tatcttatga accagtgaat caacaagaag 2220
ttaattatac tcgaccagtg atcatattgg gacctatgaa agacaggata aatgatgact 2280
tgatctcaga atttcctgac aaatttggat cctgtgttcc tcatacaact agaccaaaac 2340
gagattatga ggtagatgga agagattatc attttgtgac ttcaagagag cagatggaaa 2400
aagatatcca ggaacataaa ttcattgaag ctggccagta taacaatcat ctatatggaa 2460
caagtgttca gtctgtacga gaagtagcag aaaagggcaa acactgtatc cttgatgtgt 2520
ctggaaatgc cataaagaga ttacagattg cacagcttta ccctatctcc atttttatta 2580
aacccaaatc catggaaaat atcatggaaa tgaataagcg tctaacagaa gaacaagcca 2640
gaaaaacatt tgagagagcc atgaaactgg aacaggagtt tactgaacat ttcacagcta 2700
ttgtacaggg ggatacgctg gaagacattt acaaccaagt gaaacagatc atagaagaac 2760
aatctggttc ttacatctgg gttccggcaa aagaaaagct atgaaaactc atgtttctct 2820
gtttctcttt tccacaattc cattttcttt ggcatctctt tgccctttcc tctggagtct 2880
ttcttgagta ctgatttcat gttgaattgt atcccacaca tcatggtctg cagcttcttt 2940
ttcacatgta gtgtctcctt caagttacat cgtgtgtatt atttaatgtc actattggtt 3000
agtggccatt tttcagaact gaagatggaa tggcctgacc agctattaag aatgtgggga 3060
gacgcagaaa ttgtggtaaa attccttaat gtttaaggga aagtaacttt aagagatttt 3120
tggaaaagct ttatatacat tcttttcaaa tttcagtaca aatgaaaaag tggttttaat 3180
cagtgattta gtagactttg agcaactgtg cacgcttcag tttaatagca tggtttggcc 3240
agtgtattac tctcaagtcc ttttctcaat caacttctat catcaaagca attgtttcat 3300
tatagataaa taaggacatt ttttaattta aaaattcatg tctgagttga ctttcataag 3360
ggatttcatt ttttcctcaa cattcttaaa gccttttagt atttgacggt tctttttttcc 3420
caggacattt gctaggaata acaacgtttc aatgttttta atctacttga gcaacactat 3480
cgtgtcttac aaaagttgtt catatgtaaa tgatcatcac atttcgtgaa ttgaggccat 3540
gttcaggtgc taaggaagtt cgccttttac acagaagatt gagaaaattt cctagatata 3600
aatacagata aatcagacgt tacagtggtg acgtagtaac catcatggca atggaaagga 3660
gtccaattca tagcctaaaa cttcaaatgt attcttagga gtcagatttt actgaatatt 3720
ttacccacaa tagctgccta ttttgttata ataaaatata tataaatata tatataaaac 3780
tttctttaaa ctgtaactat gggaattatt ttctttacat agttgcgcac acacacaaac 3840
atatatatat atttaaaata tattttgttt cttttgccac ctactcctaa ctttttgttt 3900
ggtttttaaa ttaaatatta attgaataga cttatcttcc ttaatcatgt gaactgaaaa 3960
tgggggcatg gtggtcatga ggaggaaaca tttagggaaa ttagattata agtaaaagtg 4020
tggggcatat tctccccttt tctacaaagg ttttcaaatg ggtgcctgag gtttttttgtt 4080
gttgtccgtg ttgttgctgt tgtgcttt 4108

<210> 58
<211> 6050
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2658834CB1

<400> 58
ggcgggagcg gtggcggcgg cagaggcggc ggctccagct tcggctccgg ctcgggctcg 60
ggctccggct ccggctccgg ctccggctcc agctcgggtg gcggtggcgg gagcgggacc 120
aggtggaggc ggcggcggca gaggagtggg agcagcggcc ctagcggctt gcgggggggac 180
atgcggaccg acggcccctg gataggcgga aggagtggag gccctggtgc ccggcccttg 240
gtgctgagta tccagcaaga gtgaccgggg tgaagaagca aagactcggt tgattgtcct 300
gggctgtggc tggctgtgga gctagagccc tggatggccc ctgagccagc cccagggagg 360
acgatggtgc cccttgtgcc tgcactggtg atgcttggtt tggtggcagg cgcccatggt 420
gacagcaaac ctgtcttcat taaagtccct gaggaccaga ctgggctgtc aggaggggta 480
gcctccttcg tgtgccaagc tacaggagaa cccaagccgc gcatcacatg gatgaagaag 540
gggaagaaag tcagctccca gcgcttcgag gtcattgagt ttgatgatgg ggcagggtca 600

gtgcttcgga tccagccatt gcgggtgcag cgagatgaag ccatctatga gtgtacagct 660
actaacagcc tgggtgagat caacactagt gccaagctct cagtgctcga agaggaacag 720
ctgcccctg ggttcccttc catcgacatg gggcctcagc tgaaggtggt ggagaaggca 780
cgcacagcca ccatgctatg tgccgcaggc ggaaatccag accctgagat ttcttggttc 840
aaggacttcc ttcctgtaga ccctgccacg agcaacggcc gcatcaagca gctgcgttca 900
ggtgccttgc agatagagag cagtgaggaa tccgaccaag gcaagtacga gtgtgtggcg 960
accaactcgg caggcacacg ttactcagcc cctgcgaacc tgtatgtgcg agtgcgccgc 1020
gtggctcctc gtttctccat ccctcccagc agccaggagg tgatgccagg cggcagcgtg 1080
aacctgacat gcgtggcagt gggtgcaccc atgccctacg tgaagtggat gatgggggcc 1140
gaggagctca ccaaggagga tgagatgcca gttggccgca acgtcctgga gctcagcaat 1200
gtcgtacgct ctgccaacta cacctgtgtg gccatctcct cgctgggcat gatcgaggcc 1260
acagcccagg tcacagtgaa agctcttcca aagcctccga ttgatcttgt ggtgacagag 1320
acaactgcca ccagtgtcac cctcacctgg gactctggga actcggagcc tgtaacctac 1380
tatggcatcc agtaccgcgc agcgggcacg gagggcccct ttcaggaggt ggatggtgtg 1440
gccaccaccc gctacagcat tggcggcctc agccctttct cggaatatgc cttccgcgtg 1500
ctggcggtga acagcatcgg gcgagggccg cccagcgagg cagtgcgggc acgcacggga 1560
gaacaggcgc cctccagccc accgcgccgc gtgcaggcac gcatgctgag cgccagcacc 1620
atgctggtgc agtgggagcc tcccgaggag cccaacggcc tggtgcgggg ataccgcgtc 1680
tactatactc cggactcccg ccgccccccg aacgcctggc acaagcacaa caccgacgcg 1740
gggctcctca cgaccgtggg cagcctgctg cctggcatca cctacagcct gcgcgtgctt 1800
gccttcaccg ccgtgggcga tggccctccc agccccacca tccaggtcaa gacgcagcag 1860
ggagtgcctg cccagcccgc ggacttccag gccgaggtgg agtcggacac caggatccag 1920
ctctcgtggc tgctgccccc tcaggagcgg atcatcatgt atgaactggt gtactgggcg 1980
gcagaggacg aagaccaaca gcacaaggtg accttcgacc caacctcctc ctacacacta 2040
gaggacctga agcctgacac actctaccgc ttccagctgg ctgcacgctc ggatatgggg 2100
gtgggcgtct tcacccccac cattgaggcc cgcacagcac agtccatgcc cagcgggcct 2160
ccgcggaagg tggaggtgga gccactgaac tccactgctg tgcatgtcta ctggaagctg 2220
cctgtcccca gcaagcagca tggccagatc cgcggctacc aggtcaccta cgtgcggctg 2280
gagaatggcg agccccgtgg actccccatc atccaagacg tcatgctagc cgaggcccag 2340
gaaaccacta tcagcggcct gaccccggag accacctact ccgttactgt tgctgcctat 2400
accaccaagg gggatggtgc ccgcagcaag cccaaaattg tcactacaac aggtgcagtc 2460
ccaggccggc ccaccatgat gatcagcacc acggccatga acactgcgct gctccagtgg 2520
cacccaccca aggaactgcc tggcgagctg ctgggctacc ggctgcagta ctgccgggcc 2580
gacgaggcgc ggcccaacac catagatttc ggcaaggatg accagcactt cacagtcacc 2640
ggcctgcaca aggggaccac ctacatcttc cggcttgctg ccaagaaccg ggctggcttg 2700
ggtgaggagt tcgagaagga gatcaggacc cccgaggacc tgcccagcgg cttcccccaa 2760
aacctgcatg tgacaggact gaccacgtct accacagaac tggcctggga cccgccagtg 2820
ctggcggaga ggaacgggcg catcatcagc tacaccgtgg tgttccgaga catcaacagc 2880
caacaggagc tgcagaacat cacgacagac acccgcttta cccttactgg cctcaagcca 2940
gacaccactt acgacatcaa ggtccgcgca tggaccagca aaggctctgg cccactcagc 3000
cccagcatcc agtcccggac catgccggtg gagcaagtgt ttgccaagaa cttccgggtg 3060
gcggctgcaa tgaagacgtc tgtgctgctc agctgggagg ttcccgactc ctataagtca 3120
gctgtgccct ttaagattct gtacaatggg cagagtgtgg aggtggacgg gcactcgatg 3180
cggaagctga tcgcagacct gcagcccaac acagagtact cgtttgtgct gatgaaccgt 3240
ggcagcagcg caggggggcct gcagcacctg gtgtccatcc gcacagcccc cgacctcctg 3300
cctcacaagc cgctgcctgc ctctgcctac atagaggacg gccgcttcga tctctccatg 3360
ccccatgtgc aagacccctc gcttgtcagg tggttctaca ttgttgtggt acccattgac 3420
cgtgtgggcg ggagcatgct gacgccaagg tggagcacac ccgaggaact ggagctggac 3480
gagcttctag aagccatcga gcaaggcgga gaggagcagc ggcggcggcg gcggcaggca 3540
gaacgtctga agccatatgt ggctgctcaa ctggatgtgc tcccggagac ctttaccttg 3600
ggggacaaga agaactaccg gggcttctac aaccggcccc tgtctccgga cttgagctac 3660
cagtgctttg tgcttgcctc cttgaaggaa cccatggacc agaagcgcta tgcctccagc 3720
ccctactcgg atgagatcgt ggtccaggtg acaccagccc agcagcagga ggagccggag 3780
atgctgtggg tgacgggtcc cgtgctggca gtcatcctca tcatcctcat tgtcatcgcc 3840
atcctcttgt tcaaaaggaa aaggacccac tctccgtcct ctaaggatga gcagtcgatc 3900
ggactgaagg actccttgct ggcccactcc tctgaccctg tggagatgcg gaggctcaac 3960

```
taccagaccc caggtatgcg agaccaccca cccatcccca tcaccgacct ggcggacaac 4020
atcgagcgcc tcaaagccaa cgatggcctc aagttctccc aggagtatga gtccatcgac 4080
cctggacagc agttcacgtg ggagaattca aacctggagg tgaacaagcc caagaaccgc 4140
tatgcgaatg tcatcgccta cgaccactct cgagtcatcc ttacctctat cgatggcgtc 4200
cccgggagtg actacatcaa tgccaactac atcgatggct accgcaagca gaatgcctac 4260
atcgccacgc agggccccct gcccgagacc atgggtgatt tctggagaat ggtgtgggaa 4320
cagcgcacgg ccactgtggt catgatgaca cggctggagg agaagtcccg ggtaaaatgt 4380
gatcagtact ggccagcccg tggcaccgag acctgtggcc ttattcaggt gaccctgttg 4440
gacacagtgg agctggccac atacactgtg cgcaccttcg cactccacaa gagtggctcc 4500
agtgagaagc gtgagctgcg tcagtttcag ttcatggcct ggccagacca tggagttcct 4560
gagtacccaa ctcccatcct ggccttccta cgacgggtca aggcctgcaa ccccctagac 4620
gcagggccca tggtggtgca ctgcagcgcg ggcgtgggcc gcaccggctg cttcatcgtg 4680
attgatgcca tgttggagcg gatgaagcac gagaagacgg tggacatcta tggccacgtg 4740
acctgcatgc gatcacagag gaactacatg gtgcagacgg aggaccagta cgtgttcatc 4800
catgaggcgc tgctggaggc tgccacgtgc ggccacacag aggtgcctgc ccgcaacctg 4860
tatgcccaca tccagaagct gggccaagtg cctccagggg agagtgtgac cgccatggag 4920
ctcgagttca agttgctggc cagctccaag gcccacacgt cccgcttcat cagcgccaac 4980
ctgccctgca acaagttcaa gaaccggctg gtgaacatca tgccctacga attgacccgt 5040
gtgtgtctgc agcccatccg tggtgtggag ggctctgact acatcaatgc cagcttcctg 5100
gatggttata gacagcagaa ggcctacata gctacacagg ggcctctggc agagagcacc 5160
gaggacttct ggcgcatgct atgggagcac aattccacca tcatcgtcat gctgaccaag 5220
cttcgggaga tgggcaggga gaaatgccac cagtactggc cagcagagcg ctctgctcgc 5280
taccagtact ttgttgttga cccgatggct gagtacaaca tgccccagta tatcctgcgt 5340
gagttcaagg tcacggatgc ccgggatggg cagtcaagga caatccggca gttccagttc 5400
acagactggc cagagcaggg cgtgcccaag acaggcgagg gattcattga cttcatcggg 5460
caggtgcata agaccaagga gcagtttgga caggatgggc ctatcacggt gcactgcagt 5520
gctggcgtgg gccgcaccgg ggtgttcatc actctgagca tcgtcctgga gcgcatgcgc 5580
tacgagggcg tggtcgacat gtttcagacc gtgaagaccc tgcgtacaca gcgtcctgcc 5640
atggtgcaga cagaggacca gtatcagctg tgctaccgtg cggccctgga gtacctcggc 5700
agctttgacc actatgcaac gtaactaccg ctcccctctc ctccgccacc cccgccgtgg 5760
ggctccggag gggacccagc tcctctgagc cataccgacc atcgtccagc cctcctacgc 5820
agatgctgtc actggcagag cacagcccac ggggatcaca gcgtttcagg aacgttgcca 5880
caccaatcag agagcctaga acatcctggg caagtggatg gcccagcagg caggcactgt 5940
ggcccttctg tccaccagac ccacctggag cccgcttcaa gctctctgtt gcgctcccgc 6000
atttctcatg cggggatcca ctagttctag taagccgcgc cgcgtgactc            6050

<210> 59
<211> 3367
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6818489CB1

<400> 59
ggccatgatg ggctctaaag gatctggggg tggtgtggga ggattggggt cactaggagc 60
aaaatagaag gctaaggatg gagtagggag gcgaggtccc gggggcaggc gggaatgctt 120
ggaatctcgg gagtaggtag gagaactggg gtttctaaca gcgggatgga aagctgggga 180
tggggctgcg gaatgcaacc tcctacacag cctccacccc gctcccagac ccccagcact 240
gagagtgagc tctggttagg agcaaggctc tggaatcagg ccctcctggg ttcaaatcct 300
ggctccatcc catcctagct gtgcagctga atcactcagc tgctcggttt cctcatctgt 360
gatatgggga tgacaatgct cacagccacc acttcatagc aatgtcgtga ggcttccgtg 420
aattaatcta tgtcaagcgg ttggcacact gtgagtgagg ctcagagagc tctgctgctg 480
ttccttagcg attcttattt ttggtttctc tctccagcac ctggataacc acccaccttg 540
aaggagacct ccctgccctc ctgcctctgt tgtcccccag agcactgcct gatcatcctc 600
```

tgttccccat cctcccagcc cttcctgctg tacctgtggg gagctgatct cctcagtccc 660
cctgcttttc cccggtctgc catcaccacc ccaccaccat gcaccccctt cctggctact 720
ggtcctgtta ctgtctactc ctgctattct ccttgggagt ccaggggtcc ctgggggctc 780
ccagcgctgc cccagagcaa gtccatctgt cttacccagg tgagccaggc tccatgactg 840
taacttggac cacatgggtc ccaacccgct ctgaagtgca attcgggttg cagccgtcgg 900
ggcccctgcc tctccgcgcc cagggcacct tcgtcccctt tgtggacggg ggcattctcc 960
ggcggaagct ctacatacac cgagtcacgc ttcgcaagct gctgccaggg gttcagtatg 1020
tttatcgctg tggcagtgcg cagggctgga gccgtcggtt ccgcttcagg gccctcaaga 1080
atggggccca ctggagtccc cgtctggctg tgtttggaga cctgggggct gacaacccga 1140
aggccgtccc ccggctgcgc agggacaccc agcagggcat gtatgacgcc gttctccatg 1200
tgggagactt tgcctacaac ctggatcagg acaacgcccg tgttggggat aggttcatgc 1260
ggctcattga acccgtggct gccagcctgc catacatgac atgccctggg aatcatgaag 1320
aacgctacaa cttctctaac tacaaggctc gcttcagcat gccgggggat aatgagggcc 1380
tgtggtacag ctgggatctg ggtcccgccc acatcatctc cttctccacc gaggtctatt 1440
tctttctcca ttatggccgc cacttggtac agaggcagtt tcgctggctg gagagcgacc 1500
tccagaaagc caataagaac cgggcagccc ggccgtggat catcactatg ggcaccggc 1560
ccatgtactg ctccaacgca gatctggacg actgcacacg acatgaaagc aaggtccgca 1620
aaggcctcca aggcaagctg tacgggttgg aggatctttt ctacaaatat ggagtggatc 1680
tgcagctgtg ggctcatgag cactcgtatg aacgactgtg gccaatttac aactaccagg 1740
tatttaacgg cagccgagag atgccctaca ccaacccgcg agggcctgtc cacatcatca 1800
caggatctgc tggctgtgag gagcggctga cgccctttgc tgtcttcccg aggccctgga 1860
gtgccgtgcg tgtgaaggag tacgggtata cgcggctgca catcctcaac gggacccaca 1920
tccacatcca gcaggtgtcg gacgaccagg atgggaagat cgtagatgat gtctgggtgg 1980
tgagacccct gtttggccgg aggatgtacc tctagggatg gcggcagctc tcctccagaa 2040
gcctaggttt tgccgccttg gctgctgtga ccagaaactg cccaggcctg ggtggggagt 2100
tgggtgggcc ctgactcccc tgccctccag aggccccatg tagggtacat gcagccctat 2160
ggagctgggg cagctgttcc ctcctggaga ggtgggagtc ctggctggct gtggagggag 2220
ggcaggtgtg cgggcacaga gtgacacacg gcaggtttct gctggcaggg ccccaccctc 2280
ctgcatagct ctgatcgggc gaggtgccca cggggcttca ggaatgaaga ggcttaagct 2340
ctggctccat ggattctgca catctgcggg ggatgccgct gggcttcctc ctctcctgcc 2400
cacctggcaa gggcatcgcc aggtgggcac aaccgtcatg acactactca ccagcaggtg 2460
gcgtcagggg ctttttcttc tgagcccggc actgagagtt ggtctgaagc ctggctcctt 2520
cttcactgct ccaggactgc tatgaagagt cccttcatgc ctcagtttcc cagcctggca 2580
ccatcttatt cgggaagagg agacgtgtta acactcttgc ctcctagcta ggacagatga 2640
caaaccgcaa gagccacaga cttgccagtt ccttccctct ttccttcctt tctttccctt 2700
cttttattta ttgaatcata atttattgag catctaccat gtgccaggct ctgttctcag 2760
tgctggagag acagctgtga atgagacaga gatctcggcc ctcacagagc tgacatccta 2820
accagagagt tggacaaaaa tcacgataaa tgagttggtt aaatagcgat ttgtgagtag 2880
aaaacgcagg gacggtgaga gagcagtttc aattttcagg gggatctcac tgagagggca 2940
acatttgatc tgaaggaggt ggggaaggag ccaagtgggc agacatctgg gggaagagca 3000
ttccaggcag aggaacagcc agtgcaaagg ccctgagaca gaaatgtgcc tggccggctg 3060
ggtacagtga ctcacatctg tagtcccagc actttgggag gccgaggcgg gcagatcact 3120
tgagcccagg agtttgagac cagcctgggc aacacagcga aaccctttct ctaccggaaa 3180
tatgaaattt agctgggcat ggtggtgtgt gcctgtggtc ccggctgctc gggaggctga 3240
ggcgggagga tggcttgggc ccaggaggtt gaggctgcag tgagccatga ttgcgccact 3300
gcaccccagc ctgggcaaca gagcaagacc ctgtctcaaa aaatatatat aaataaacaa 3360
agtttta 3367

<210> 60
<211> 2512
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509415CB1

<400> 60
ggtcttccct tttcactaag cctctgttgc acagctccat ggacattttg tggcacaggg 60
actcccctaa gggggagcca aagcccaggc ctcccctaca agggaaaggg tcccagtctg 120
gtatgccccc ttggtcagca agtcttcctg gcatcagtta gagtatgtct ctctcccctc 180
ggcctgttca ctgtcatttt cctcttcttc attggtcaac atgtgtgtcc tgttgctgat 240
tttaaacatc tctagtttgt agtctgtatc tcccttgcac tggcctttat acaatctttc 300
ccctggtaac tcctggagca tctgtccagt ggggccagga agaatgcaga gcaggagtgg 360
agctcctctc ctccatatct gccaggtggt cctgggtccg cagcccgccc tcacaggccc 420
tcctcactcc cctaggtaga tggcccctc agggcaggcc cggcggacac ccctccctct 480
ggctggcgga tgcagtgcct agcggccgcc cttaaggacg aaaccaacat gagtggggga 540
ggggagcagg ccgacatcct gccggccaac tacgtggtca aggatcgctg gaaggtgctg 600
aaaaagatcg gggcggggg ctttggtgag atctacgagg ccatggacct gctgaccagg 660
gagaatgtgg ccctcaaggt ggagtcagcc cagcagccca agcaggtcct caagatggag 720
gtggccgtgc tcaagaagtt gcaagggaag gaccatgtgt gcaggttcat tggctgtggc 780
aggaacgaga agtttaacta tgtagtgatg cagctccagg gccggaacct ggccgacctg 840
cgccgtagcc agccgcgagg caccttcacg ctgagcacca cattgcggct gggcaagcag 900
accttggagt ccatcgaggc catccactct gtgggcttcc tgcaccgtga catcaagcct 960
tcaaactttg ccatgggcag gctgccctcc acctacagga agtgctatat gctggacttc 1020
gggctggccc ggcagtacac caacaccacg ggggatgtgc ggccccctcg gaatgtggcc 1080
gggtttcgag gaacggttcg ctatgcctca gtcaatgccc acaagaaccg ggagatgggc 1140
cgccacgacg acctgtggtc cctcttctac atgctggtgg agtttgcagt gggccagctg 1200
ccctggagga agatcaagga caaggaacag gtagggatga tcaaggagaa gtatgagcac 1260
cggatgctgc tgaagcacat gccgtcagag ttccacctct tcctggacca cattgccagc 1320
ctcgactact tcaccaagcc cgactaccag ttgatcatgt cagtgtttga gaacagcatg 1380
aaggagaggg gcattgccga gaatgaggcc tttgactggg agaaggcagg caccgatgcc 1440
ctcctgtcca cgagcacctc taccccgccc cagcagaaca cccggcagac ggcagccatg 1500
tttggggtgg tcaatgtgac gccagtgcct ggggacctgc tccgggagaa caccgaggat 1560
gtgctacagg gagagcacct gagtgaccag gagaatgcac ccccaattct gcccgggagg 1620
ccctctgagg ggctgggcca cagtccccac cttgtccccc accccggggg tcctgaggct 1680
gaagtctggg aggagacaga tgtcaaccgg aacaaactcc ggatcaacat cggcaaagta 1740
actgccgcca gggcgaaggg cgtgggtggc cttttctctc acccccgatt cccagccttg 1800
tgcccctgcc ctgttcctcc taagcaccct gtccccggcc atctgcctgc ttgccctgcc 1860
tctgtttccc ggtccctccc cgcactagcc tcgctgtgtc ttccatcatc atcatcctct 1920
gtctccttca ccctgaggag accatccgcc cacagccgcc tcatcagccc cagctcatgg 1980
cactcccctc tcctgcagag cccctgtgtg gaggaggaac agagccgagg catgggggtc 2040
cccagctccc cagtgcgtgc ccccccagac tcccccacaa ccccagtccg ttctctgcgc 2100
taccggaggg tgaacagccc tgagtcagaa aggctgtcca cggcggacgg gcgagtggag 2160
ctacctgaga ggaggtgggt ctggggccag gggcatggtt ggggcccaag gccctctccg 2220
ccttcacgtg gctggtctgg aggaaaagtt agatgtgtgg cggaggtggg cagaccctgg 2280
gaagtgctga gagggttata cttgggcctg gggtcagact cagttggggc cagagacagg 2340
gcctgggaga accagtgggg gatccagaga ggtcccggct catgccagga aacgtaattg 2400
ggtgagtgca ggctgcagga gggacaggtg gggcgcctgg gcccaggaag ggtgaggggc 2460
gagttggttg gtgggtgtgt gtgcttccag aatctcttct cctagagact aa 2512

<210> 61
<211> 2147
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506916CB1

<220>
<221> unsure
<222> (1) ... (2147)

```
<223> a, t, c, g, or other

<400> 61
tgccccgccc ccggctcggc cccgccccta cccttcccga agctccgcac ctgccacgcc 60
ccggcgctca gctccgcctc gacccttcc aggagctccg ccccgccccc gccccttcca 120
gaagccccgc ccctgccccg cccccagcgc tcgcaatcta cggggcgggt tccaccgccc 180
agaagagcgc acggcgggcg ctgccggccc cgagcgctcc cgctaccact gccgctcccg 240
gccctgccgg ccggccatgc cccgagttcc ggctgccgcg cgcacccaga cccggcgatg 300
aggagtggcg ccgagcgcag gggcagcagc gccgcggcgt ccccgggctc gccgcccccc 360
ggccgcgcgc gccccgccgg ctccgacgcg ccctcggccc tgccgccgcc cgctgctggc 420
cagccccggg cccgggactc gggcgatgtc cgctcgcagc cgcgccccct gtttcagtgg 480
agcaagtgga agaagaggat gggctcgtcc atgtcggcgg ccaccgcgcg gaggccggtg 540
tttgacgaca aggaggacgt gaacttcgac cacttccaga tccttcgggc cattgggaag 600
ggcagctttg gcaaggtgtg cattgtgcag aagcgggaca cggagaagat gtacgccatg 660
aagtacatga acaagcagca gtgcatcgag cgcgacgagg tccgcaacgt cttccgggag 720
ctggagatcc tgcaggagat cgagcacgtc ttcctggtga acctctggta ctccttccag 780
gacgaggagg acatgttcat ggtcgtggac ctgctactgg gcggggacct gcgctaccac 840
ctgcagcaga acgtgcagtt ctccgaggac acggtgaggc tgtacatctg cgagatggca 900
ctggctctgg actacctgcg cggccagcac atcatccaca gagatgtcaa gcctgacaac 960
attctcctgg atgagagagg acatgcacac ctgaccgact tcaacattgc caccatcatc 1020
aaggacgggg agcgggcgac ggcattagca ggcaccaagc cgtacatggc tccggagatc 1080
ttccactctt ttgtcaacgg cgggaccggc tactccttcg aggtggactg gtggtcggtg 1140
ggggtgatgg cctatgagct gctgcgagga tggaggccct atgacatcca ctccagcaac 1200
gccgtggagt ccctggtgca gctgttcagc accgtgagcg tccagtatgt ccccacgtgg 1260
tccaaggaga tggtggcctt gctgcggaag aaaggccgtc tgcactgcga ccccaccttt 1320
gagctggagg agatgatcct ggagtccagg cccctgcaca agaagaagaa gcgcctggcc 1380
aagaacaagt cccgggacaa cagcagggac agctcccagt ccgagaatga ctatcttcaa 1440
gactgcctcg atgccatcca gcaagacttc gtgattttta acagagaaaa gctgaagagg 1500
agccaggacc tcccgaggga gcctctcccc gcccctgagt ccagggatgc tgcggagcct 1560
gtggaggacg aggcggaacg ctccgccctg cccatgtgcg gccccatttg cccctcggcc 1620
gggagcggct aggccgggac gcccgtggtc ctcacccctt gagctgcttt ggagactcgg 1680
ctgccagagg gagggccatg ggccgaggcc tggcattcac gttcccaccc agcctggctg 1740
gcggtgccca cagtgccccg gacacatttc acacctcagn nnnnnnnnnn nnnnnnnnnn 1800
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1860
nnnnnnnnnn gcgcgctcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg 1920
cgttacccaa cttaatcgcc ttgcagcaca tccccctttt cggcaaggct ggcgtaatag 1980
gcgaagagtg cccgcaccga tcgcccttcc caacagttgc gcagctgaat gggcgaatgg 2040
gacgcgccct gtagcggcgc ttaagcgcgg cgggtgtggg ggtacgcgcg cgtgaccggt 2100
tccttgcgcg gccctagccg cgtcttgctc ttcccttctt tttcccg              2147

<210> 62
<211> 1421
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507104CB1

<400> 62
ccaccacctc ccttagcctc ctgtgtggga ggagtttatg ggtatgtggc tcctgcccag 60
tccaggtggg ctttcacttc tactctattt cagttcctct ttcccgatct gggctggaga 120
gcttcctcat tgttaaggca gcagaaactt tcgctggatg gttttaggat aaggggtcat 180
caatgctggc aagagtcggc acaatgagga ccaggcttgc tgtgaagtgg tgtatgtgga 240
aggtcggagg agtgttacag gagtacctag ggagcctagc cgaggccagg gactctgctt 300
ctactactgg ggcctatttg atgggcatgc agggggcgga gctgctgaaa tggcctcacg 360
```

gctcctgcat cgccatatcc gagagcagct aaaggacctg gtagagatac ttcaggaccc 420
ttcgccacca cccctctgcc tcccaaccac tccggggacc ccagattcct ccgatccctc 480
tcacttgctt ggccctcagt cctgctggtc ttcacagaag gaagtgagcc acgagagcct 540
ggtagtgggg gccattgaga atgccttcca gctcatgggc catcattgtc cggaatggtg 600
aaatcattcc aatgtcccgg gagtttaccc cggagactga gcgccagcgt cttcagctgc 660
ttggcttcct gaaaccagag ctgctaggca gtgaattcac ccaccttgag ttcccccgca 720
gagttctgcc caaggagctg gggcagagga tgttgtaccg ggaccagaac atgaccggct 780
gggcctacaa aaagatcgag ctggaggatc tcaggtttcc tctggtctgt ggggagggca 840
aaaaggctcg ggtgatggcc accattgggg tgacccgagg cttgggagac cacagcctta 900
aggtctgcag ttccaccctg cccatcaagc cctttctctc ctgcttccct gaggtacgag 960
tgtatgacct gacacaatat gagcactgcc cagatgatgt gctagtcctg ggaacagatg 1020
gcctgtggga tgtcactact gactgtgagg tagctgccac tgtggacagg gtgctgtcgg 1080
cctatgagcc taatgaccac agcaggtata cagctctggc ccaagctctg gtcctggggg 1140
cccggggtac cccccgagac cgtggctggc gtctccccaa caacaagctg ggttccgggg 1200
atgacatctc tgtcttcgtc atccccctgg gagggccagg cagttactcc tgaggggctg 1260
aacaccatcc ctcccactag cctctccata cttactcctc tcacagccca aattctgaag 1320
ttgtctccct gacccttctt tagtggcaac ttaactgaag aagggatgtc cgctatatcc 1380
aaaattacag ctattggcaa ataaacgaga tggataaaaa a 1421

<210> 63
<211> 1481
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507105CB1

<400> 63
ccaccacctc ccttagcctc ctgtgtggga ggagtttatg ggtatgtggc tcctgcccag 60
tccaggtggg ctttcacttc tactctattt cagttcctct ttcccgatct gggctggaga 120
gcttcctcat tgttaaggca gcagaaactt tcgctggatg gttttaggat aaggggtcat 180
caatgctggc aagagtcggc acaatgagga ccaggcttgc tgtgaagtgg tgtatgtgga 240
aggtcggagg agtgttacag gagtacctag ggagcctagc cgaggccagg gactctgctt 300
ctactactgg ggcctatttg atgggcatgc agggggcgga gctgctgaaa tggcctcacg 360
gctcctgcat cgccatatcc gagagcagct aaaggacctg gtagagatac ttcaggaccc 420
ttcgccacca cccctctgcc tcccaaccac tccggggacc ccagattcct ccgatccctc 480
tcacttgctt ggccctcagt cctgctggtc ttcacagaag gaagatgagc agatggcccg 540
ggagcggcgt ggccaccaag tggagggggg ctgctgtgca ctggttgtga tctacctgct 600
aggcaaggtg tacgtggcca atgcaggcga tagcagggcc atcattgtcc ggaatggtga 660
aatcattcca atgtcccggg agtttacccc ggagactgag cgccagcgtc ttcagctgct 720
tggcttcctg aaaccagagc tgctaggcag tgaattcacc caccttgagt tcccccgcag 780
agttctgccc aaggagctgg ggcagaggat gttgtaccgg gaccagaaca tgaccggctg 840
ggcctacaaa aagatcgagc tggaggatct caggtttcct ctggtctgtg gggagggcaa 900
aaaggctcgg gtgatggcca ccattggggt gacccgaggc ttgggagacc acagccttaa 960
ggtctgcagt tccaccctgc ccatcaagcc ctttctctcc tgcttccctg aggtacgagt 1020
gtatgacctg acacaatatg agcactgccc agatgatgtg ctagtcctgg gaacagatgg 1080
cctgtgggat gtcactactg actgtgaggt agctgccact gtggacaggg tgctgtcggc 1140
ctatgagcct aatgaccaca gcaggtatac agctctggcc caagctctgg tcctgggggc 1200
ccggggtacc ccccgagacc gtggctggcg tctccccaac aacaagctgg gttccgggga 1260
tgacatctct gtcttcgtca tccccctggg agggccaggc agttactcct gaggggctga 1320
acaccatccc tcccactagc ctctccatac ttactcctct cacagcccaa attctgaagt 1380
tgtctccctg acccttcttt agtggcaact taactgaaga agggatgtcc gctatatcca 1440
aaattacagc tattggcaaa taaacgagat ggataaaaaa a 1481

<210> 64

<211> 1472
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507107CB1

<400> 64
ccaccacctc ccttagcctc ctgtgtggga ggagtttatg ggtatgtggc tcctgcccag 60
tccaggtggg ctttcacttc tactctattt cagttcctct ttcccgatct gggctggaga 120
gcttcctcat tgttaaggca gcagaaactt tcgctggatg gttttaggat aaggggtcat 180
caatgctggc aagagtcggc acaatgagga ccaggcttgc tgtgaagtgg tgtatgtgga 240
aggtcggagg agtgttacag gagtacctag ggagcctagc cgaggccagg gactctgctt 300
ctactactgg ggcctatttg atgggcatgc agggggcgga gctgctgaaa tggcctcacg 360
gctcctgcat cgccatatcc gagagcagct aaaggacctg gtagagatac ttcaggaccc 420
ttcgccacca cccctctgcc tcccaaccac tccggggacc ccagattcct ccgatccctc 480
tcacttgctt ggccctcagt cctgctggtc ttcacagaag gaagtgagcc acgagagcct 540
ggtagtgggg gccattgaga atgccttcca gctcatggat gagcagatgg cccgggagcg 600
gcgtggccac caagtggagg ggggctgctg tgcactggtt gtgatctacc tgctaggcaa 660
ggtgtacgtg gccaatgcag gcgatagcag ggccatcatt gtccggaatg gtgaaatcat 720
tccaatgtcc cgggagttta ccccggagac tgagcgccag cgtcttcagc tgcttggctt 780
cctgaaacca gagctgctag gcagtgaatt cacccacctt gagttccccc gcagagttct 840
gcccaaggag ctggggcaga ggatgttgta ccgggaccag aacatgaccg gctggctcgg 900
gtgatggcca ccattggggt gacccgaggc ttgggagacc acagccttaa ggtctgcagt 960
tccaccctgc ccatcaagcc ctttctctcc tgcttccctg aggtacgagt gtatgacctg 1020
acacaatatg agcactgccc agatgatgtg ctagtcctgg gaacagatgg cctgtgggat 1080
gtcactactg actgtgaggt agctgccact gtggacaggg tgctgtcggc ctatgagcct 1140
aatgaccaca gcaggtatac agctctggcc caagctctgg tcctgggggc ccggggtacc 1200
ccccgagacc gtggctggcg tctccccaac aacaagctgg gttccgggga tgacatctct 1260
gtcttcgtca tccccctggg agggccaggc agttactcct gaggggctga acaccatccc 1320
tcccactagc ctctccatac ttactcctct cacagcccaa attctgaagt tgtctccctg 1380
acccttcttt agtggcaact taactgaaga agggatgtcc gctatatcca aaattacagc 1440
tattggcaaa taaacgagat ggataaaaaa aa 1472

<210> 65
<211> 1480
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507109CB1

<400> 65
ccaccacctc ccttagcctc ctgtgtggga ggagtttatg ggtatgtggc tcctgcccag 60
tccaggtggg ctttcacttc tactctattt cagttcctct ttcccgatct gggctggaga 120
gcttcctcat tgttaaggca gcagaaactt tcgctggatg gttttaggat aaggggtcat 180
caatgctggc aagagtcggc acaatgagga ccaggcttgc tgtgaagtgg tgtatgtgga 240
aggtcggagg agtgttacag gagtacctag ggagcctagc cgaggccagg gactctgctt 300
ctactactgg ggcctatttg atgggcatgc agggggcgga gctgctgaaa tggcctcacg 360
gctcctgcat cgccatatcc gagagcagct aaaggacctg gtagagatac ttcaggaccc 420
ttcgccacca cccctctgcc tcccaaccac tccggggacc ccagattcct ccgatccctc 480
tcacttgctt ggccctcagt cctgctggtc ttcacagaag gaagtgagcc acgagagcct 540
ggtagtgggg gccattgaga atgccttcca gctcatggat gagcagatgg cccgggagcg 600
gcgtggccac caagtggagg ggggctgctg tgcactggtt gtgatctacc tgctaggcaa 660

ggtgtacgtg gccaatgcag gcgatagcag ggccatcatt gtccggaatg gtgaaatcat 720
tccaatgtcc cgggagttta ccccggagac tgagcgccag cgtcttcagc tgcttggctt 780
cctgaaacca gagctgctag gcagtgaatt cacccacctt gagttccccc gcagagttct 840
gcccaaggag ctggggcaga ggatgttgta ccgggaccag aacatgaccg gctgggccta 900
caaaaagatc gagctggagg atctcaggtt tcctctggtc tgtggggagg gcaaaaaggc 960
tcgggtgatg gccaccattg gggtgacccg aggcttggga gaccacagcc ttaaggtctg 1020
cagttccacc ctgcccatca agccctttct ctcctgcttc cctgaggtac gagtgtatga 1080
cctgacacaa tatgagcact gcccagatga tgtgctagtc ctgggaacag atggcctgtg 1140
ggatgtcact actgactgtg aggtatacag ctctggccca agctctggtc ctgggggccc 1200
ggggtacccc ccgagaccgt ggctggcgtc tccccaacaa caagctgggt tccggggatg 1260
acatctctgt cttcgtcatc cccctgggag ggccaggcag ttactcctga ggggctgaac 1320
accatccctc ccactagcct ctccatactt actcctctca cagcccaaat tctgaagttg 1380
tctccctgac ccttctttag tggcaactta actgaagaag ggatgtccgc tatatccaaa 1440
attacagcta ttggcaaata aacgagatgg ataaaaaaaa 1480

<210> 66
<211> 515
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1833937CB1

<400> 66
cgctgtgtga atcgtgggtg ggatggccgc gggccgcctc tttctaagtc ggcttcgagc 60
acccttcagt tccatggcca agagcccact cgagggcgtt tcctcctcca gaggcctgca 120
cgcggggcgc gggccccgaa ggctctccat cgaaggcaac attggctcca ctttgaggct 180
ctgatgaaca ttccagtgct ggtgttggat gtcaatgatg atttttctga ggaagtaacc 240
aaacaagaag acctcatgag agaggtaaac acctttgtaa agaatctgta accaatacca 300
tgatgttcag gctgtgatct gggctccctg actttctgaa gctagaaaaa tgttgtgtct 360
cccaaccacc tttccatccc cagcccctct catccctgga gcactctgcc gctcaagagc 420
tggtttgtta attattgtta gactttgcca ttgttttctt ttgtacctga agcattttga 480
aaataaagtt tacttaagtt taaaaaaaaa aaaaa 515

<210> 67
<211> 1961
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7502036CB1

<400> 67
gtgcaccgac gtcggcgcgg gctgcaccgc cgcgtccgcc cgcccgccag catggccacc 60
accgccacct gcacccgttt caccgacgac taccagctct tcgaggagct tggcaagggt 120
gctttctctg tggtccgcag gtgtgtgaag aaaacctcca cgcaggagta cgcagcaaaa 180
atcatcaata ccaagaaatt gtctgcccgg gatcaccaga aactagaacg tgaggctcgg 240
atatgtcgac ttctgaaaca tccaaacatc gtgttaccgg cggggagctg tttgaagaca 300
ttgtggccag agagtactac agtgaagcag atgccagcca ctgtatacat cagattctgg 360
agagtgttaa ccacatccac cagcatgaca tcgtccacag ggacctgaag cctgagaacc 420
tgctgctggc gagtaaatgc aagggtgccg ccgtcaagct ggctgatttt ggcctagcca 480
tcgaagtaca gggagagcag caggcttggt ttggttttgc tggcaccccca ggttacttgt 540
cccctgaggt cttgaggaaa gatccctatg gaaaacctgt ggatatctgg gcctgcgggg 600
tcatcctgta tatcctcctg gtgggctatc ctcccttctg ggatgaggat cagcacaagc 660

```
tgtatcagca gatcaaggct ggagcctatg atttcccatc accagaatgg gacacggtaa 720
ctcctgaagc caagaacttg atcaaccaga tgctgaccat aaacccagca aagcgcatca 780
cggctgacca ggctctcaag cacccgtggg tctgtcaacg atccacggtg gcatccatga 840
tgcatcgtca ggagactgtg gagtgtttgc gcaagttcaa tgcccggaga aaactgaagg 900
gtgccatcct cacgaccatg cttgtctcca ggaacttctc agttggcagg cagagctccg 960
cccccgcctc gcctgccgcg agcgccgccg gcctggccgg gcaagctgcc aaaagcctat 1020
tgaacaagaa gtcggatggc ggtgtcaagc cacagagcaa caacaaaaac agtctcgtaa 1080
gcccagccca agagcccgcg cccttgcaga cggccatgga gccacaaacc actgtggtac 1140
acaacgctac agatgggatc aagggctcca cagagagctg caacaccacc acagaagatg 1200
aggacctcaa agctgccccg ctccgcactg ggaatggcag ctcggtgcct gaaggacgga 1260
gctcccggga cagaacagcc ccctctgcag gcatgcagcc ccagccttct ctctgctcct 1320
cagccatgcg aaaacaggag atcattaaga ttacagaaca gctgattgaa gccatcaaca 1380
atggggactt tgaggcctac acgaagattt gtgatccagg cctcacttcc tttgagcctg 1440
aggcccttgg taacctcgtg gaggggatgg atttccataa gttttacttt gagaatctcc 1500
tgtccaagaa cagcaagcct atccatacca ccatcctaaa cccacacgtc cacgtgattg 1560
gggaggacgc agcgtgcatc gcctacatcc gcctcaccca gtacatcgac gggcagggtc 1620
ggcctcgcac cagccagtca gaagagaccc gggtctggca ccgtcgggat ggcaagtggc 1680
tcaatgtcca ctatcactgc tcagggggccc ctgccgcacc gctgcagtga gctcagccac 1740
aggggcttta ggagattcca gccggaggtc caaccttcgc agccagtggc tctggagggc 1800
ctgagtgaca gcggcagtcc tgtttgtttg aggtttaaac catttcattt ccaaaggggg 1860
aacagccatt gaaggccctt gaatgaagcc ttccggcccg cctttggggt tgtcttttgt 1920
tttaccaggt gtttttaca atttaggaaa aaaaaaaaa a                     1961
```

```
<210> 68
<211> 1154
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503248CB1

<400> 68
ggctgcgccg agagcggaga cacaggctca agatggcaga ttccgactga ggctgggggg 60
gccgagctcg cgcgccgctt tcccgtcccc gttgccatga accgcggaca ccccggcccc 120
gatggccccc gtgtacgaag gtatggcctc acatgtgcaa gttttctccc ctcacaccct 180
tcaatcaagt gccttctgta gtgtgaagaa actgaaaata gagccgagtt ccaactggga 240
catgactggg tacggctccc acagcaaagt gtatagccag agcaagaaca tccccctgtc 300
gcagccagcc accacaaccg tcagcacctc cttgccggtc ccaaacccaa gcctacctta 360
cgagcagacc atcgtcttcc caggaagcac cgggcacatc gtggtcacct cagcaagcag 420
cacttctgtc accgggcaag tcctcggcgg accacacaac ctaatgcgtc gaagcactgt 480
gagcctcctt gatacctacc aaaaatgtgg actcaagcgt aagagcgagg agatcgagaa 540
cacaagcagc gtgcagatca tcgaggagca tccacccatg attcagaata atgcaagcgg 600
ggccactgtc gccactgcca ccacgtctac tgccacctcc aaaaacagcg gctcgcccac 660
catccacccg agtcagtatc cagcccaatt tgcccaccag acctacatca gcgcctcgcc 720
agcctccacc gtctacactg gatacccact gagccccgcc aaggtcaacc agtaccctta 780
catataaaca ctggagggga gggagggagg gagggaggga gagaatggcc cgagggagga 840
gggagagaag gagggaggcg ctcctgggac cgtgggcgct ggccttttat actgaagatg 900
ccgcacacaa acaatgcaaa cggggcaggt gcgggggggg ggggggcaga gggcaggggc 960
acggggtcgg gacaccagtg aaacttgaac cgggaagtgg gaggacgtag agcagagaag 1020
agaacatttt taaaggaag ggattaaaga gggtgggaaa tctatggttt ttattttaaa 1080
aaagaaaaag gaaaaaaaaa aagtcaataa caaaaaaccc agctcaagaa cccattctac 1140
gtcaaactgg aaat                                                  1154

<210> 69
<211> 1938
```

```
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503968CB1

<400> 69
gcggtccctc ccggcccggc ggaacgcgtc ccttttaagg gggcggggac ctgggggtct 60
ggggccagcg cgcgggaggg acgcctgagt gcctcgaggg cgccgttcgg gcggggagga 120
tcccgcgggt cccactgacc cacgcggggt ggggccaggg gtggacgctc gcccgtacgc 180
ggtcgctact gatcatgctt gggccagggt ccaatcgcag gcgccccacg caggggggagc 240
gaggcccagg gtcccccgga gagcccatgg agaagtacca ggttttgtac cagctgaatc 300
ctggggcctt gggggtgaac ctggtggtgg aggaaatgga aaccaaagtc aagcatgtga 360
taaagcaggt ggaatgcatg gatgaccatt acgccagtca ggccctggag gagctgatgc 420
cactgctgaa gctgcggcac gcccacatct ctgtgtacca ggagctgttc atcacgtgga 480
atggggagat ctcttctctg tacctctgcc tggtgatgga gttcaatgag ctcagcttcc 540
aggaggtcat tgaggataag aggaaggcaa agaaaatcat tgactctgag tggatgcaga 600
atgtgctggg ccaggtgctg gacgcgctgg aatacctgca ccatttggac atcatccaca 660
ggaatctcaa accctccaac atcatcctca tcagcagtga ccactgcaaa ctgcaggacc 720
tgagttccaa tgtgctaatg acagacaaag ccaaatggaa tattcgtgcg gaggaagacc 780
cctttcgtaa gtcctggatg gcccctgaag ccctcaactt ctccttcagc cagaaatcag 840
acatctggtc cctgggctgc atcattctgg acatgaccag ctgctccttc atggatggca 900
cagaagccat gcatctgcgg aagtccctcc gccagagccc aggcagcctg aaggccgtcc 960
tgaagacaat ggaggagaag cagatcccgg atgtggaaac cttcaggaat cttctgccct 1020
tgatgctcca gatcgacccc tcggatcgaa taacgataaa gtctgccgtg gcccccggag 1080
ctggtggagg tggtggtcac gaccatggag ctacatgaca gggtcctcga tgtccagctg 1140
tgtgcctgct ccctgctgct gcacctcctg ggccaaggta tcattgtgaa caaggccccc 1200
ttggagaagg tcccggacct catcagccag gtgttggcca cctaccctgc ggatggggaa 1260
atggcagaag ccagctgcgg agtcttctgg ctgctgtccc tgctgggctg catcaaggag 1320
cagcagtttg aacaagtggt ggcgctgctc ctgcaaagca tccggctgtg ccaggacaga 1380
gccctgctgg tgaacaatgc ctaccgggga ctggccagcc tggtgaaggt gtcagagctg 1440
gcggccttca aggtggtggt gcaggaggag ggcggcagtg gcctcagcct catcaaggag 1500
acctaccagc tccacaggga cgacccggag gtggtggaga acgtgggcat gctgctggtc 1560
cacctggctt cctatgagga gatcctgccg gagctggtgt ccagtagtat gaaggccctg 1620
ctccaggaga tcaaggagcg cttcacctcc agcctggaac tggtttcttg cgcggaaaaa 1680
gtgctcttga ggctggaggc agccacctct cccagcccac tgggtgggga agcagctcag 1740
ccctgatgcg ggggagaaga cagatacccc acaggcccct ccctccacgt gtgccctctc 1800
cctgtccttc ctttccatgg gccactgttt cccttggggt ggggggaagg gtcatccagc 1860
accagaatgc gcacctcaca ctcctcttag gtgactaata aagaggccca aggccagttt 1920
ctgccttaaa aaaaaagg                                                1938

<210> 70
<211> 931
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505931CB1

<400> 70
gcgcggcctc gatctctgga ctcgtcacct gcccctcccc ctcccgccgc cgtcacccag 60
gaaaccggcc gcaatcgccg gccgacctga agctggtttc atggcagcct caaagaaggc 120
agttttgggg ccattggtgg gggcggtgga ccagggcacc agttcgacgc gctttttggt 180
tttcaattca aaaacagctg aactacttag tcatcatcaa gtagaaataa aacaagagtt 240
```

cccaagagaa ggatgggtgg aacaggaccc taaggaaatt ctacattctg tctatgagtg 300
tatagagaaa acatgtgaga aacttggaca gctcaatatt gatatttcca acataaaagc 360
tattggtgtc agcaaccaga gggaaaccac tgtagtctgg gacaagataa ctggagagcc 420
tctctacaat gctgtggaaa gtgaaattcg ttattctaca tggaagaaag ctgtgatgaa 480
gtcaatgggt tgggttacaa ctcaatctcc agaaagtggt attccataaa acctaccaac 540
tcatggattc ccaagatgtg agctttttac ataatgaaag aacccagcaa ttctgtctct 600
taatgcaatg acactattca tagactttga ttttatttat aagccacttg ctgcatgacc 660
ctccaagtag acctgtggct taaaataaag aaaatgcagc aaaaagaatg ctatagaaat 720
atttggtggt tttttttttt tttaaacatc cacagttaag gttgggccag ctacctttgg 780
ggctgacccc ctccattgcc ataacatcct gctccattcc ctctaagatg taggaagaat 840
tcggatcctt accattggaa tcttccatcg acatactcaa cactattgga ccaggattga 900
gtctctgcat gcataacttg agtagagggg g 931

<210> 71
<211> 1730
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506912CB1

<400> 71
atgaagctta taaatggcaa aaagcaaaca agaaggcctt caccagatgc agtccctcga 60
tcttggactt ctcagcctcc agaactcatt cccatggttt ggcatggaca tcggtggaac 120
gctggttaaa ttggtgtatt tcgagccgaa ggatattaca gccgaagagg agcaagagga 180
agtggagaac ctgaagagca tccggaagta tttgacttct aatactgctt atgggaaaac 240
tgggatccga gacgtccacc tggaactgaa aaacctgacc atgtgtggac gcaaagggaa 300
cctgcacttc atccgctttc ccagctgtgc tatgcacagg ttcattcaga tgggcagcga 360
gaagaacttc tctagccttc acaccaccct ctgtgccaca ggaggcgggg ctttcaaatt 420
cgaagaggac ttcagaatga ttgctgacct gcagctgcat aaactggatg aactggactg 480
tctgattcag ggcctgcttt atgtcgactc tgttggcttc aacggcaagc cagaatgtta 540
ctattttgaa aatcccacaa atcctgaatt gtgtcaaaaa aagccgtact gccttgataa 600
cccatacccct atgttgctgg ttaacatggg ctcaggtgtc agcattctag ccgtgtactc 660
caaggacaac tataaaagag ttacagggac cagtcttgga ggtggaacat tcctaggcct 720
atgttgcttg ctgactggtt gtgagacctt tgaagaagct ctggaaatgg cagctaaagg 780
cgacagcacc aatgttgata aactggtgaa ggacatttac ggaggagact atgaacgatt 840
tggccttcaa ggatctgctg tagcatcaag ctttggcaac atgatgagta aagaaaagcg 900
agattccatc agcaaggaag acctcgcccg ggccacattg gtcaccatca ccaacaacat 960
tggctccatt gctcggatgt gcgcgttgaa tgagaacata gacagagttg tgtttgttgg 1020
aaatttctc agaatcaata tggtctccat gaagctgctg gcatatgcca tggattttg 1080
gtccaaagga caactgaaag ctctgtttt ggaacatgag ggttattttg gagccgttgg 1140
ggcactgttg gaactgttca aaatgactga tgataagtag agacgagcag tggaggaaac 1200
agcctcccaa aaggacagag aactaaaaaa ttgctgctgg agaaggtgaa agtcgctttg 1260
ggacggaagc caagccatta tggcagatga acctgctgga tttgtaaata atttaaaatc 1320
cttccagatg atcttttact cttaggtttt gagctaatga ttcaaaacgg gggaatataa 1380
aaggtttttt ttctgtatac tgtatttttt taaaaaaatg gtgcagcgtg gccaaaccta 1440
ccaattgtat gcattaactt tgaaaagttg tttgatgttt aagaaggacc tgatatgtaa 1500
gcgctggtca tttttcttct ggggtttact gatcagtgtg gtgattttaa cttcatttag 1560
taattactct aggagatttt accttgactt atatttttca tgacgtttca tgatttgctg 1620
ttggtttcaa atgaaactac aaatctggca tgttttactg tgaacacttt tgttatttgt 1680
tttgtaccct tttttgtctt gtttttctgt tttagttgtc ttctgaaaaa 1730

<210> 72
<211> 1495
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506913CB1

<400> 72
atgaagctta taaatggcaa aaagcaaaca ttcccatggt ttggcatgga catcggtgga 60
acgctggtta aattggtgta tttcgagccg aaggatatta cagccgaaga ggagcaagag 120
gaagtggaga acctgaagag catccggaag tatttgactt ctaatactgc ttatgggaaa 180
actgggatcc gagacgtcca cctggaactg aaaaacctga ccatgtgtgg acgcaaaggg 240
aacctgcact tcatccgctt tcccagctgt gctatgcaca ggttcattca gatgggcagc 300
gagaagaact tctctagcct tcacaccacc ctctgtgcca caggaggcgg ggctttcaaa 360
ttcgaagagg acttcagaat gattgctgac ctgcagctgc ataaactgga tgaactggac 420
tgtctgattc agggcctgct ttatgtcgac tctgttggct tcaacggcaa gccagaatgt 480
tactattttg aaaatcccac aaatcctgaa ttgtgtcaaa aaaagccgta ctgccttgat 540
aacccatacc ctatgttgct ggttaacatg ggctcaggtg tcagcattct agccgtgtac 600
tccaaggaca actataaaag agttacaggg accagctttg gcaacatgat gagtaaagaa 660
aagcgagatt ccatcagcaa ggaagacctc gcccgggcca cattggtcac catcaccaac 720
aacattggct ccattgctcg gatgtgcgcg ttgaatgaga acatagacag agttgtgttt 780
gttggaaatt ttctcagaat caatatggtc tccatgaagc tgctggcata tgccatggat 840
ttttggtcca aaggacaact gaaagctctg tttttggaac atgagggtta ttttggagcc 900
gttggggcac tgttggaact gttcaaaatg actgatgaca agtagagacg agcagtggag 960
gaaacagcct cccaaaagga cagagaacta aaaaattgct gctggagaag gtgaaagtcg 1020
ctttgggacg gaagccaagc cattatggca gatgaacctg ctggatttgt aaataattta 1080
aaatccttcc agatgatctt ttactcttag gttttgagct aatgattcaa aacgggggaa 1140
tataaaaggt tttttttctg tatactgtat ttttttaaaa aaatggtgca gcgtggccaa 1200
acctaccaat tgtatgcatt aactttgaaa agttgtttga tgtttaagaa ggacctgata 1260
tgtaagcgct ggtcattttt cttctggggt ttactgatca gtgtggtgat tttaacttca 1320
tttagtaatt actctaggag attttacctt gacttatatt tttcatgacg tttcatgatt 1380
tgctgttggt ttcaaatgaa actacaaatc tggcatgttt tactgtgaac acttttgtta 1440
tttgttttgt accctttttt gtcttgtttt tctgttttag ttgtcttctg aaaaa 1495

<210> 73
<211> 1153
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507029CB1

<400> 73
caattttttt tttttttttt tttttttct ccaacgaagt cttatatttt accattctgc 60
gtgtattaga aagctaaatt taaacagcaa caataagaaa tcttccccag tggatccctc 120
gagcctttcg ggaaaatgca ggtttctggc tgacaggaag aagaaacagc atccaatcgg 180
cacttgtgtc tcaatgacca atcaaatttc gccttacatt gcatcgctgg gataaacgga 240
gctggacgac tcagtctctt ggtctgtggc tgctgcggtt acctggatgg gcgagcacct 300
ctgaggctgg ctttgttacc tgggcaataa gggactagca gttcagccgt tttctatgcc 360
tgctggattt gtttgtattt gttcccagcc actgctcatg taatgtactc ccttaaccag 420
gaaattaaag cattctcccg gaataatctc aggaagcaat gcaccagggt gacaacgcta 480
actggaaaga aaattataga aacatggaaa gatgccagaa ttcatgttgt ggaagaagta 540
gagccgagca gtgggggtgg ttgtggttat gtgcaggacc ttagctcgga cctgcaagtt 600
ggcgttatta agccatggtt gctcctaggg tcacaagatg ctgctcatga tttggataca 660
ctgaaaaaga ataaggatgg agtggttctt gttcattgta atgcaggcgt ttccagggct 720
gctgcaattg taataggttt cctgatgaat tctgaacaaa cctcatttac cagtgctttt 780

tctttggtga aaaatgcaag accttccata tgtccaaatt ctggcttcat ggagcagctt 840
cgtacatatc aagagggcaa agaaagcaat aagtgtgaca gaatacagga gaacagttca 900
tgagttgcat tgtagcagac aatggacaac tgtagtttct gaattgactt ctatagccat 960
cttttccctt ttttggagag tagactagca aaactccctt ttttctcttg cctttttat 1020
gcataaatgg aggtcaattt gattgtcctg acctactgta taagtaaatt tcaaatgtca 1080
ttactttctc tttgttatta taatgtgtga ttaaatgctt ttttaaattg ctaagggaaa 1140
ataaaaaaaa aaa 1153

<210> 74
<211> 1472
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507063CB1

<400> 74
gctggaggca ggggccatgc cctgcaagag cgccgagtgg ctgcaggagg agctggaggc 60
gcgcggcggc gcgtccttgc tgctgctcga ctgccggccg cacgagctct tcgagtcgtc 120
gcacatcgag acggccatca acctggccat cccgggcctc atgttgcgcc gcctgcgcaa 180
gggcaacctg cccatccgct ccatcatccc caaccacgcc gacaaggagc gcttcgccac 240
gcgctgcaag gcggccaccg tgctgctcta cgacgaggcc acggccgagt ggcagcccga 300
gcccggcgct cccgcctccg tgctcggcct gctcctacag aagctgcgcg acgacggctg 360
ccaggcctac tacctccaag atcctgccct acctctacct cggctgcgcc aaggactcca 420
ccaacctgga cgtgctcggc aagtatggca tcaagtatat cctcaatgtc acacccaacc 480
tacccaacgc cttcgagcac ggcggcgagt tcacctacaa gcagatcccc atctctgacc 540
actggagcca gaacctctcc cagttcttcc ctgaggccat cagcttcatt gacgaagccc 600
gctccaagaa gtgtggtgtc ctggtgcact gcctggcagg catcagccgc tcagtgacgg 660
tcactgtggc ctatctgatg cagaagatga acctgtcact caacgacgcc tacgactttg 720
tcaagaggaa aaagtccaac atctcgccca acttcaactt catggggcag ctgctggact 780
ttgagcggac gctggggcta agcagcccgt gcgacaacca cgcgtcgagt gagcagctct 840
acttttccac gcccaccaac cacaacctgt tcccactcaa tacgctggag tccacgtgag 900
gcctggtgca cgggggcat ggcaccaggc ccctgctcgg ctctccacag ggctaggtgg 960
gagagcccaa gcccgccacc tctggcctga ggaaccccca gatgtcacct gtgcccagag 1020
gcccaggctg atcggtgtcg gagcgcccct caccatcctt gggggcaggg cccgcaggca 1080
aggtctccca ctgcagggct tgctggagag gcctcggctc ttggacacgt ggctttgggc 1140
gtccaccagg gcctcatcct gtccaggacg ctcctttctg ctgacagccc agccagtttg 1200
gctgtttttt aaagacacat ccacggacct gagtttactt tttacttttg gcaggtaaat 1260
ccaagctccc tggagcacaa agagtgtttg agctcttctt gatttttctt tttttttttt 1320
tttttaaca aaaagtgtta ttttcaggct acatgcaaca gtggattgta taacccagta 1380
tttcatccct ttcctgatcc gtgcaagaga gagaaatgtt cagttttcaa ctttaatcat 1440
tgtgaattac cttatgcgat tttaagaact gg 1472

<210> 75
<211> 2997
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504755CB1

<400> 75
gccagggctg ccgtggcaaa atgagctgtc aactttaggt tgacaggggt gtggccgcga 60
ccgcaagggc ttttgttgcc gggtggaccc aacagggatg ggctgctggg gacagctgct 120

ggtgtggttc ggagccgcgg gcgccattct ctgctctagc ccggggtccc aggagacttt 180
tctgcggtcc tcgcccctgc cgctggcaag tcccagcccc caggacccga aagtcagcgc 240
cccgcctagt atcttggagc cagcctcccc gctgaattct ccgggcaccg aggggtcttg 300
gctgttttct acctgcgggg ccagcggccg gcatgggccc acacagacac aatgtgacgg 360
ggcgtacgcg gggaccagcg tggtggtgac cgtgggggcc gccgggcagc tgagaggcgt 420
gcagctgtgg cgcgtgccgg gccctggcca gtatctgatc tcagcctacg gagccgcggg 480
cggcaaaggc gccaagaacc acctgtcgcg ggcgcatggc gtcttcgtct cagcaatctt 540
ctccctcggt ctcggggagt cgctgtacat cctggtgggg cagcagggag aggacgcctg 600
tcccggaggt agcccggaga gccagctcgt ctgcctcggg gagtctcgag ccgttgaaga 660
gcacgcggcg atggatggga gcgaaggggt cccggggtcg cggcgctggg cgggaggtgg 720
cgggggtggc gggggcgcca cctacgtttt ccgggtgcgc gctggcgagc tggaaccgtt 780
gctggtggcg gccggaggcg gcggtcgggc ctacctgagg ccgcgggacc gaggccggac 840
tcaggcctcc cccgagaaac tggagaaccg ctcggaggcg cccgggagcg gcgggagagg 900
cggggcggca ggtggtgggg gcggctggac gtcgcgggct ccctctccgc aggccggccg 960
ctcactgcag gagggggcgg agggcggcca gggctgctcc gaggcttggg cgacccttgg 1020
ctgggccgcg gccggcggct tcgggggcgg cggcggggcc tgcactgcgg gcggaggcgg 1080
cggcggctac agggggggcg acgcttcaga gactgacaac ctctgggctg atggggaaga 1140
tggagtatcc ttcatacacc ccagcagcga gctcttcctg cagcctctgg cagtcaccga 1200
gaaccacgga gaggtagaga tccgaaggca cctcaactgc agtcactgcc ctttgagaga 1260
ctgccaatgg caggcagagc tccagctggc tgaatgcctg tgcccagaag gcatggagct 1320
agctgtggat aacgtcacct gcatggacct gcacaagccc ccaggccctc tggttctgat 1380
ggtggctgtg gtggcaacct caacactgag cctccttatg gtgtgtgggg tcctgattct 1440
gggtggggct tggcccggcc cagtcctggc ctctgccacc aggtgtcacc gaggtttccc 1500
cagccaatgt tactctgctc agagccctgg gccatggtgc ctttggggag gtgtatgagg 1560
gactggtaat tggccttcct ggggactcca gtcccctgca ggtagctatc aagaccctgc 1620
cagaactctg ctcgcctcag gatgagctgg atttcctcat ggaggccctc atcatcagca 1680
agtttcgcca tcagaacatt gtgcggtgtg tggggctcag cctcagggcc acccctcgcc 1740
tcattctgct ggaactgatg tctggagggg acatgaagag tttcctgagg cacagtcggc 1800
cacacctggg ccagccatca cctctggtca tgcgggacct gctgcaactg gcccaggaca 1860
tagcccaggg ctgccactac ctggaggaaa atcacttcat ccacagggat attgccgccc 1920
ggaactgcct gctgagctgc gctggaccca gccgagtggc caagattggg gactttggga 1980
tggcacgaga tatctaccgg gccagttatt accgcagggg ggaccgggcc ttgctcccag 2040
tcaagtggat gcccccagag gccttcctgg agggcatctt cacatccaag acagattcct 2100
ggtcttttgg ggtgctgctc tgggagatct tctcactggg ctacatgccc tatcctgggc 2160
gcaccaacca ggaggtgctg gacttcgtcg ttggaggagg ccggatggac cctcctaggg 2220
gctgcccagg gcctgtgtac cgcatcatga cccagtgttg gcagcacgag cctgagctcc 2280
gccctagctt tgccagcatc ttggagcgtc tgcagtactg cactcaggac ccggatgtgc 2340
tgaattcact cctgccaatg gagctggggc ccaccccaga ggaggaaggg acttctgggc 2400
tggggaacag atctttggag tgcctaagac cccacagcc ccaggaactg agtccagaga 2460
agttgaaaag ctggggaggt agccctcttg gcccctggct gtcctctggc ctcaagcccc 2520
tcaaatccag ggcctccaa cctcagaacc tttggaatcc cacttatcgc tcctgagccc 2580
caaggggccc tgagggtaag gactgaggca ctgagggtcc ctccctatac tcctcaggct 2640
cctgggtggc ctgttatgcc agcggcctct gttccctgca gtctgtgctg tgtgtctggg 2700
cctgtctcgg ggctggcctg gcagcgctgc acttgccatg ctggaaacca gcccaggcct 2760
cccagggaag ggcccagcca cttccagctt ttgatcttgg ggccagaggc cgccttacac 2820
acaccccagg tgtccatggg gagcactgga ttgctctccc attatgagca tccttcatct 2880
gggcagaccc cccatcctgc agatgcttct aataaaagc tcttctcatc ctccaaaaaa 2940
aaaaaagggc gcgctcagag atcagctacg taggtgatgc gagtatagtt tcaaggt 2997

<210> 76
<211> 2046
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7509265CB1

<400> 76
gctgtcatcg ttccgtgggc cctgctgcgg gcacgctctc ggcgcatgcg tttttatgc 60
gggattaagc ttgctgctgc gtgacagcgg agggctagga aaaggcgcag tggggcccgg 120
agctgtcacc cctgactcga cgcagcttcc gttctcctgg tgacgtcgcc tacaggaacc 180
gccccagtgg tcagctgccg cgctgttgct aggcaacagc gtgcgagctc agatcagcgt 240
ggggtggagg agaagtggag tttggaagtt caggggcaca ggggcacagg cccacgactg 300
cagcgggatg gaccagtact gcatcctggg ccgcatcggg gagggcgccc acggcatcgt 360
cttcaaggcc aagcacgtgg agccgagggt gggctggcag tgtctgcctt ctatcctgca 420
gactggcgag atagttgccc tcaagaaggt ggccctaagg cggttggaag acggcttccc 480
taaccaggcc ctgcgggaga ttaaggctct gcaggagatg gaggacaatc agtatgtggt 540
acaactgaag gctgtgttcc cacacggtgg aggctttgtg ctggcctttg agttcatgct 600
gtcggatctg gccgaggtgg tgcgccatgc ccagaggcca ctagcccagg cacaggtcaa 660
gagctacctg cagatgctgc tcaagggtgt cgccttctgc catgccaaca acattgtaca 720
tcgggacctg aaacctgcca acctgctcat cagcgcctca ggccagctca agatagcgga 780
ctttggcctg gctcgagtct tttccccaga cggcagccgc ctctacacac accaggtggc 840
caccaggagc tcactgagct gccggactac aacaagatct cctttaagga gcaggtgccc 900
atgcccctgg aggaggtgct gcctgacgtc tctccccagg cattggatct gctgggtcaa 960
ttccttctct accctcctca ccagcgcatc gcagcttcca aggctctcct ccatcagtac 1020
ttcttcacag ctcccctgcc tgcccatcca tctgagctgc cgattcctca gcgtctaggg 1080
ggacctgccc ccaaggccca tccagggccc ccccacatcc atgacttcca cgtggaccgg 1140
cctcttgagg agtcgctgtt gaacccagag ctgattcggc ccttcatcct ggaggggtga 1200
gaagttggcc ctggtcccgt ctgcctgctc ctcaggacca ctcagtccac ctgttcctct 1260
gccacctgcc tggcttcacc ctccaaggcc tccccatggc cacagtgggc ccacaccaca 1320
ccctgcccct tagcccttgc gagggttggt ctcgaggcag aggtcatgtt cccagccaag 1380
agtatgagaa catccagtcg agcagaggag attcatggcc tgtgctcggt gagccttacc 1440
ttctgtgtgc tactgacgta cccatcagga cagtgagctc tgctgccagt caaggcctgc 1500
atatgcagaa tgacgatgcc tgccttggtg ctgcttcccc gagtgctgcc tcctggtcaa 1560
ggagaagtgc agagagtaag gtgtccttat gttggaaact caagtggaag gaagatttgg 1620
tttggtttta ttctcagagc cattaaacac tagttcagta tgtgagatat agattctaaa 1680
aacctcaggt ggctctgcct tatgtctgtt cctccttcat ttctctcaag ggaaatggct 1740
aaggtggcat tgtctcatgg ctctcgtttt tggggtcatg ggagggtag caccaggcat 1800
agccactttt gccctgaggg actcctgtgt gcttcacatc actgagcact catttagaag 1860
tgagggagac agaagtctag gcccagggat ggctccagtt ggggatccag caggagaccc 1920
tctgcacatg aggctggttt accaacatct actccctcag gatgagcgtg agccagaagc 1980
agctgtgtat ttaaggaaac aagcgttcct ggaattaatt tataaattta ataaatccca 2040
attatt 2046

<210> 77
<211> 1668
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509371CB1

<400> 77
gctcatttcg gcgaaaccgc ggtctttcct tctccccttg atgctttcag gtactgaccc 60
actaccgccc ccatcttccc ccatgggaag atgagcactg agggcagatt accctcctgc 120
agcgcgtgtg tgaaagggga gttgagagtg ctgacgagcg cggcgctcac tagtcgggac 180
ggcccgagac cgtgtcatgt cctcttcagg attgtgcacc tgtgcctgcg aaaggctgac 240
cagaagctgg tgatcatcaa gcagattcca gtggaacaga tgaccaagga agagcggcag 300
gcagcccaga atgagtgcca ggtcctcaag ctgctcaacc accccaatgt cattgagtac 360
tacgagaact tcctggaaga caaagccctt atgatcgcca tggaatatgc accaggcggc 420

actctggctg agttcatcca aaagcgctgt aattccctgc tggaggagga gaccatcctg 480
cacttcttcg tgcagatcct gcttgcactg catcatgtgc acacccacct catcctgcac 540
cgagacctca agacccagaa catcctgctt gacaaacacc gcatggtcgt caagatcggt 600
gatttcggca tctccaagat ccttagcagc aagagcaagg cctacacggt ggtgggtacc 660
ccatgctata tctcccctga gctgtgtgag ggcaagccct acaaccagaa gagtgacatc 720
tgggccctgg gctgtgtcct ctacgagctg gccagcctca agagggcttt cgaggctgcg 780
aacttgccag cactggtgct gaagatcatg agtggcacct ttgcacctat ctctgaccgg 840
tacagccctg agcttcgcca gctggtcctg agtctactca gcctggagcc tgcccagcgg 900
ccaccactca gccacatcat ggcacagccc ctctgcatcc gtgccctcct caacctccac 960
accgacgtgg gcagtgtccg catgcggagg gcagagaagt ccgtggcccc cagcaacaca 1020
gggagcagga ccaccagtgt ccgctgcaga ggtaagtggg aagaggccgc cagtccccat 1080
ggatgccaca ccattcccat cagtttaata gtccccatgc actgtacctg caggtatccc 1140
ccggggacct gtgaggccag ccatcccacc accactgtcg tcagtgtatg cctggggtgg 1200
tgggctgggc accccccctgc ggctgccaat gctcaacaca gaggtggtcc aggtggcagc 1260
tgggcgcacg cagaaagccg gcgtcacgcg ctctgggcgt ctcatcctgt gggaggcccc 1320
acccctaggt gcaggcggag gcagtctcct tcctggggca gtggagcagc cacagcccca 1380
gttcatctcg cgtttcctgg agggccagtc gggtgtgacc atcaagcacg tggcctgtgg 1440
ggacttcttc actgcctgcc tgactgacag aggcatcatc atgacattcg gcagcggcag 1500
caatgggtgc ctaggccatg gcagcctcac tgacatcagc cagcccacca ttgtggaggc 1560
tttgctgggc tatgaaatgg tgcaggtggc ctgtggggcc tctcacgtgc tggccctgtc 1620
cactgagcga gaactatttg cctggggccg tggagacagc ggtaagct 1668

<210> 78
<211> 1735
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509389CB1

<400> 78
gctcatttcg gcgaaaccgc ggtctttcct tctccccttg atgctttcag gtactgaccc 60
actaccgccc ccatcttccc ccatgggaag atgagcactg agggcagatt accctcctgc 120
agcgcgtgtg tgaaagggga gttgagagtg ctgacgagcg cggcgctcac tagtcgggac 180
ggcccgagac cgtgtcatgt cctcttcagg attgtgcacc tgtgcctgcg aaaggctgac 240
cagaagctgg tgatcatcaa gcagattcca gtggaacaga tgaccaagga agagcggcag 300
gcagcccaga atgagtgcca ggtcctcaag ctgctcaacc accccaatgt cattgagtac 360
tacgagaact tcctggaaga caaagccctt atgatcgcca tggaatatgc accaggcggc 420
actctggctg agttcatcca aaagcgctgt aattccctgc tggaggagga gaccatcctg 480
cacttcttcg tgcagatcct gcttgcactg catcatgtgc acacccacct catcctgcac 540
cgagacctca agacccagaa catcctgctt gacaaacacc gcatggtcgt caagatcggt 600
ggtgggtacc ccatgctata tctcccctga gctgtgtgag ggcaagccct acaaccagaa 660
gagtgacatc tgggccctgg gctgtgtcct ctacgagctg gccagcctca agagggcttt 720
cgaggctgcg aacttgccag cactggtgct gaagatcatg agtggcacct ttgcacctat 780
ctctgaccgg tacagccctg agcttcgcca gctggtcctg agtctactca gcctggagcc 840
tgcccagcgg ccaccactca gccacatcat ggcacagccc ctctgcatcc gtgccctcct 900
caacctccac accgacgtgg gcagtgtccg catgcggagg cctgtgcagg gacagcgagc 960
ggtcctgggc ggcagggtgt gggcacccag tgggagcaca ggaggtctga ggcagaggga 1020
aacctggggc aagtcctccc ttcctgcatg taggaatgtc aggagggtct ttgtccttag 1080
gcccccatct gtcctgcagg gcagagaagt ccgtggcccc cagcaacaca gggagcagga 1140
ccaccagtgt ccgctgcaga ggtatccccc ggggacctgt gaggccagcc atcccaccac 1200
cactgtcgtc agtgtatgcc tggggtggtg ggctgggcac ccccctgcgg ctgccaatgc 1260
tcaacacaga ggtggtccag gtggcagctg ggcgcacgca gaaagccggc gtcacgcgct 1320
ctgggcgtct catcctgtgg gaggccccac ccctaggtgc aggcggaggc agtctccttc 1380
ctggggcagt ggagcagcca cagccccagt tcatctcgcg tttcctggag ggccagtcgg 1440

gtgtgaccat caagcacgtg gcctgtgggg acttcttcac tgcctgcctg actgacagag 1500
gcatcatcat gacattcggc agcggcagca atgggtgcct aggccatggc agcctcactg 1560
acatcagcca gcccaccatt gtggaggctt tgctgggcta tgaaatggtg caggtggcct 1620
gtggggcctc tcacgtgctg gccctgtcca ctgagcgaga actatttgcc tggggccgtg 1680
gagacaacgg caaacttggg gctaaggacc agggaattcc caaaagtggc ccaag 1735

<210> 79
<211> 1413
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7507005CB1

<400> 79
cgctgctgca gcagccgcag cgccggccgc ggctccggct ccggctccgg ctcccgggca 60
tttaaagggg acgcggcggc tgcccggggg ggatgagggg caagtggagg ggacggctca 120
gacgcacatc atcctcagtc cctcgggact ggagggactc gtgagccgga gcccagaaat 180
ccgggggtgg ataagacacc gcgtcccctc caattcccgt aagcacccct tgctccatcc 240
tgcgccccaa tacctcagct agccccttc cccacttctt acactccaaa ctcagccggg 300
acagacctct gctgccgccg cccccacgaa cgtgtgacga cggctggagg ccaacagagt 360
ccctacaggt ggtgctcacg gtaatgcacc gacaatgagt ggctgttttc cagtttctgg 420
cctccgctgc ctatctaggg tgttggatgt gccagcgcct ctggccgtgg cccgctaacc 480
agcctctccc gggcgggctc ctgccgcgcc ccctctcgct tgccccctcc tcctcctcct 540
cctgctgctc tccccctgc tcccaggacg gcaggatggc cgcgcagggc gcgccgcgct 600
tcctcctgac cttcgacttc gacgagacta tcgtggacga aaacagcgac gattcgatcg 660
tgcgcgccgc gccgggccag cggctcccgg agagcctgcg agccacctac cgcgagggct 720
tctacaacga gtacatgcag cgcgtcttca agtacctggg cgagcagggc gtgcggccgc 780
gggacctgag cgccatctac gaagccatcc ctttgtcgcc aggcatgagc gacctgctgc 840
agtttgtggc aaaacagggc gcctgcttcg aggtgattct catctccgat gccaacacct 900
ttggcgtgga gagctcgctg cgcgccgccg gccaccacag cctgttccgc cgcatcctca 960
gcaacccgtc ggggccggat gcgcggggac tgctggctct gcggccgttc cacacacaca 1020
gctgcgcgcg ctgccccgcc aacatgtgca agcacaaggt gctcagcgac tacctgcgcg 1080
agcgggccca cgacggcgtg cacttcgagc gcctcttcta cgtgggcgac ggcgccaacg 1140
acttctgccc catggggctg ctggcgggcg gcgacgtggc cttcccgcgc cgcggctacc 1200
ccatgcaccg cctcattcag gaggcccaga aggccgagcc cagctcgttc cgcgccagcg 1260
tggtgccctg ggaaacggct gcagatgtgc gcctccacct gcaacaggtg ctgaagtcgt 1320
gctgagtctg gccgcctgca ggggggtacc cgggccaacg gcggaggggg cggggaaggg 1380
agattcggca aagacagctt tactactccc tta 1413

<210> 80
<211> 935
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509142CB1

<400> 80
cgtggaccca actggcgagg ctgctggggt tgcagcggga cagttggggc ggccccgcag 60
gcccaggtga acaaaaattg tttgctggcc cccaggatac taactagacc tttggcctga 120
ctcacaggac actaaggctc cttttctgaa gaagcctttt accagtctca tttaggggat 180
gggaacaaca tgtcttcacc tagggacttt agagcagagc ctgtaaacga ttatgaggga 240
aatgactctg aagcagaaga cttgaatttc agggagactt tgccttcatc aagtcaggaa 300

aacacaccta gatcaaaggt ttttgaaaat aaagttaatt cagagaaggt aaaactttct 360
cttcggaatt tcccacataa tgattatgag gatgtttttg aagagccttc agaaagtggc 420
agtgatccca gcatgtggac agccagaggc cccttcagaa gagacaggtg gagcagtgag 480
gatgaggagg ctgcagggcc atcacaggct ctctcccctc tactttctga tacgcgcaaa 540
attgtttctg aaggagaact agatcagttg gctcagattc ggccattaat attcaatttt 600
catgagcaga cagccatcaa ggattgtttg aaaatccttg aggaaaaaac agcagcgtat 660
gatatcatgc aggaatttat ggctttagaa cttaagaatc tgcctggtga gttcaactct 720
gggaatcaac caagcaacag agaaaaaaac agataccgag atattcttcc atttcaacat 780
catggatata gtggcccaaa tgagagaaca acgttctggc atggttcaaa cgaaggagca 840
gtatcacttt tgttacgata ttgtgcttga agttcttcgg aaacttctga ctttggatta 900
agaaagactt ctgttgcctc tcacttgaaa ttacc 935

<210> 81
<211> 2038
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509157CB1

<400> 81
aattcggctc gagcgtggac ccaactggcg aggctgctgg ggttgcagcg ggacagttgg 60
ggcggccccg caggcccagg tgaacaaaaa ttgtttgctg gcccccagga tactaactag 120
acctttggcc tgactcacag gacactaagg ctccttttct gaagaagcct tttaccagtc 180
tcatttaggg gatgggaaca acatgtcttc acctagggac tttagagcag agcctgtgtt 240
ttctatttga tcagttttc cccccgcctt gttcagtaaa cgattatgag ggaaatgact 300
ctgaagcaga agacttgaat ttcagggaga ctttgccttc atcaagtcag gaaaacacac 360
ctagatcaaa ggtttttgaa aataaagtta attcagagaa ggtaaaactt tctcttcgga 420
atctcccaca taatgattat gaggatgttt ttgaagagcc ttcagaaagt ggcagtgatc 480
ccagcatgtg gacagccaga ggccccttca gaagagacag gtggagcagt gaggatgagg 540
aggctgcagg gccatcacag gctctctccc ctctactttc tgtattccat tcaccagttg 600
aaggacactg ggttgtttcc agatttggca attatggata aaagtactac agatattcac 660
aaactttgtg taaacataag tttccattac tttaagatac gcgcaaaatt gtttctgaag 720
gagaactaga tcagttggct cagattcggc cattaatatt caattttcat gagcagacag 780
ccatcaagga ttgtttgaaa atccttgagg aaaaaacagc agcgtatgat atcatgcagg 840
aatttatggc tttagaactt aagaatctgc ctggtgagtt caactctggg aatcaaccaa 900
gcaacagaga aaaaaacaga taccgagata ttcttccata tgattcaaca cgcgttcctc 960
ttggaaaaag caaggactac atcaatgcta gttatattag aatagtcaat tgtggagaag 1020
agtattttta tatcgctact caaggaccac tgctgagcac catagatgac ttttggcaaa 1080
tggtgttgga aaataattca aatgttattg ccatgataac cagagagata gaaggtggaa 1140
ttatcaaatg ctaccattac tggcccattt ctctgaagaa gccattggaa ttgaaacact 1200
tccgtgtatt cctggagaac taccagatac ttcaatattt catcattcga atgtttcaag 1260
ttgtggagaa gtccacggga actagtcact ctgtaaaaca gttgcagttc accaagtggc 1320
cagaccatgg cactcctgcc tcagcagata gcttcataaa atatattcgt tatgcaagga 1380
agagccacct tacaggaccc atggttgttc actgcagtgc cggcataggc cggacagggg 1440
tgttcctatg tgtggatgtc gtgttctgtg ccatcgtaaa gaactgttca ttcaacatca 1500
tggatatagt ggcccaaatg agagaacaac gttctggcat ggttcaaacg aaggagcagt 1560
atcacttttg ttacgatatt gtgcttgaag ttcttcggaa acttctgact ttggattaag 1620
aaagacttct gttgcctctc acttgaaatt accaagtggg tttgcacctc ctcataaaga 1680
acatgtttgc actgtgctga agggctttgc tatgcataca atctgctttc ttggtttatc 1740
agtttatttt ctttctaaaa gctccctgaa gggcaatatc atttggcttg ggtgatcag 1800
tgtttactta ttgatcttgc tagacaatat caaaataact tcccacattt tccagtgaaa 1860
cagatgttac ataaaacgat tgcagcttgg ctatttggtt gaagggattt acagagccca 1920
ataaaggatt taaaatatat tcattaagat tttatttgga aaggtggctg gagagagctg 1980
aggatttcca ggactttgta agttcttatt ctgggagaac ataaggccaa taatcatg 2038

<210> 82
<211> 1471
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509246CB1

<400> 82
gctggaggca ggggccatgc cctgcaagag cgccgagtgg ctgcaggagg agctggaggc 60
gcgcggcggc gcgtccttgc tgctgctcga ctgccggccg cacgagctct tcgagtcgtc 120
gcacatcgag acggccatca acctggccat cccgggcctc atgttgcgcc gcctgcgcaa 180
gggcaacctg cccatccgct ccatcatccc caaccacgcc gacaaggagc gcttcgccac 240
gcgctgcaag gcggccaccg tgctgctcta cgacgaggcc acggccgagt ggcagcccga 300
gcccggcgct cccgcctccg tgctcggcct gctcctacag aagctgcgcg acgacggctg 360
ccaggcctac tacctccaag atcctgccct acctctacct cggctgcgcc aaggactcca 420
ccaacctgga cgtgctcggc aagtatggca tcaagtatat cctcaatgtc acacccaacc 480
tacccaacgc cttcgagcac ggcggcgagt tcacctacaa gcagatcccc atctctgacc 540
actggagcca gaacctctcc cagttcttcc ctgaggccat cagcttcatt gacgaagccc 600
gctccaagaa gtgtggtgtc ctggtgcact gcctggcagg catcagccgc tcagtgacgg 660
tcactgtggc ctatctgatg cagaagatga acctgtcact caacgacgcc tacgactttg 720
tcaagaggaa aaagtccaac atctcgccca acttcaactt catggggcag ctgctggact 780
ttgagcggac gctggggcta agcagcccgt gcgacaacca cgcgtcgagt gagcagctct 840
acttttccac gcccaccaac cacaacctgt tcccactcaa tacgctggag tccacgtgag 900
gcctggtgca cgggggggcat ggcaccaggc ccctgctcgg ctctccacag ggctaggtgg 960
gagagcccaa gcccgccacc tctggcctga ggaacccccca gatgtcacct gtgcccagag 1020
gcccaggctg atcggtgtcg gagcgcccct caccatcctt gggggcaggg cccgcaggca 1080
aggtctccca ctgcagggct tgctggagag gcctcggctc ttggacacgt ggctttgggc 1140
gtccaccagg gcctcatcct gtccaggacg ctcctttctg ctgacagccc agccagtttg 1200
gctgtttttt aaagacacat ccacggacct gagtttactt tttacttttg gcaggtaaat 1260
ccaagctccc tggagcacaa agagtgtttg agctcttctt gatttttctt tttttttttt 1320
ttttttaaca aaaagtgtta ttttcaggct acatgcaaca gtggattgta taacccagta 1380
tttcatccct ttcctgatcc tgcaagagag agaaatgttc agttttcaac tttaatcatt 1440
gtgaattacc ttatgcgatt ttaagaactg g 1471

<210> 83
<211> 1601
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509380CB1

<400> 83
aattcggctc gagcgtggac ccaactggcg aggctgctgg ggttgcagcg ggacagttgg 60
ggcggccccg caggcccagg tgaacaaaaa ttgtttgctg gcccccagga tactaactag 120
acctttggcc tgactcacag gacactaagg ctcctttttct gaagaagcct tttaccagtc 180
tcatttaggg gatgggaaca acatgtcttc acctagggac tttagagcag agcctgtaaa 240
cgattatgag ggaaatgact ctgaagcaga agacttgaat ttcagggaga ctttgccttc 300
atcaagtcag gaaaacacac ctagatcaaa ggtttttgaa aataaagtta attcagagaa 360
ggtaaaactt tctcttcgga atttcccaca taatgattat gaggatgttt ttgaagagcc 420
ttcagaaagt ggcagtgatc ccagcatgtg gacagccaga ggccccttca gaagagacag 480
gtggagcagt gaggatgagg aggctgcagg gccatcacag gctctctccc ctctactttc 540
tgatacgcgc aaaattgttt ctgaaggaga actagatcag ttggctcaga ttcggccatt 600

aatattcaat tttcatgagc agacagccat caaggattgt ttgaaaatcc ttgaggaaaa 660
aacagcagcg tatgatatca tgcaggaatt tatggcttta gaacttaaga atctgcctgg 720
tgagttcaac tctgggaatc aaccaagcaa cagagaaaaa aacagatacc gagatattct 780
tccatatgat tcaacacgcg ttcctcttgg aaaaagcaag gactacatca atgctagtta 840
tattagaata gtcaattgtg gagaagagta tttttatatc gctactcaag gaccactgct 900
gagcaccata gatgactttt ggcaaatggt gttggaaaat aattcaaatg ttattgccat 960
gataaccaga gagatagaag gtggaattat caaatgctac cattactggc ccatttctct 1020
gaagaagcca ttggaattga aacacttccg tgtattcctg gagaactacc agatacttca 1080
atatttcatc attcgaatgt ttcaagttgt ggagaagtcc acgggaacta gtcactctgt 1140
aaaacagttg cagttcacca agtggccaga ccatggcact cctgcctcag cagatagctt 1200
cataaaatat attcgttatg caaggaagag ccaccttaca ggacccatgg ttgttcactg 1260
cagtgccggc ataggccgga caggggtgtt cctatgtgtg gatgtcgtgt tctgtgccat 1320
cgtaaagaac tgttcattca acatcatgga tatagtggcc caaatgagag aacaacgttc 1380
tggcatggtt caaacgaagg aactttccag tgtggctgaa aattacaaat gtgttttctg 1440
aaagtggatt ctgttattca attacttaac attctgttgt tttttatata tatgtatata 1500
ggagcagtat cacttttgtt acgatattgt gcttgaagtt cttcggaaac ttctgacttt 1560
ggattaagaa agacttctgt tgcctctcac ttgaaattac c 1601

<210> 84
<211> 1404
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509382CB1

<400> 84
aattcggctc gagcgtggac ccaactggcg aggctgctgg ggttgcagcg ggacagttgg 60
ggcggccccg caggcccagg tgaacaaaaa ttgtttgctg gcccccagga tactaactag 120
acctttggcc tgactcacag gacactaagg ctccttttct gaagaagcct tttaccagtc 180
tcatttaggg gatgggaaca acatgtcttc acctagggac tttagagcag agcctgtaaa 240
cgattatgag ggaaatgact ctgaagcaga agacttgaat ttcagggaga ctttgccttc 300
atcaagtcag gaaaacacac ctagatcaaa ggtttttgaa aataaagtta attcagagaa 360
ggtaaaactt tctcttcgga atttcccaca taatgattat gaggatgttt ttgaagagcc 420
ttcagaaagt ggcagtgatc ccagcatgtg gacagccaga ggccccttca gaagagacag 480
gtggagcagt gaggatgagg aggctgcagg gccatcacag gctctctccc ctctactttc 540
tgatacgcgc aaaattgttt ctgaaggaga actagatcag ttggctcaga ttcggccatt 600
aatattcaat tttcatgagc agacagccat caaggattgt ttgaaaatcc ttgaggaaaa 660
aacagcagcg tatgatatca tgcaggaatt tatggcttta gaacttaaga atctgcctgg 720
tgagttcaac tctgggaatc aaccaagcaa cagagaaaaa aacagatacc gagatattct 780
tccatacggg aactagtcac tctgtaaaac agttgcagtt caccaagtgg ccagaccatg 840
gcactcctgc ctcagcagat agcttcataa aatatattcg ttatgcaagg aagagccacc 900
ttacaggacc catggttgtt cactgcagtg ccggcatagg ccggacaggg gtgttcctat 960
gtgtggatgt cgtgttctgt gccatcgtaa agaactgttc attcaacatc atggatatag 1020
tggcccaaat gagagaacaa cgttctggca tggttcaaac gaaggagcag tatcactttt 1080
gttacgatat tgtgcttgaa gttcttcgga aacttctgac tttggattaa gaaagacttc 1140
tgttgcctct cacttgaaat taccaagtgg gtttgcacct cctcataaag aacatgtttg 1200
cactgtgctg aagggctttg ctatgcatac aatctgcttt cttggtttat cagtttattt 1260
tctttctaaa agctccctga agggcaatat catttggctt ggggtgatca gtgtttactt 1320
attgatcttg ctagacaata tcaaaataac ttcccacatt ttccagtgaa acagatgtta 1380
cataaaacga ttgcagcttg gctg 1404